Polarized CMB power spectrum estimation using the pure pseudo cross-spectrum approach

J. Grain,[1, 2, *] M. Tristram,[3, †] and R. Stompor[1, ‡]

[1]CNRS, Laboratoire AstroParticule & Cosmologie, Université Paris 7, Denis Diderot
10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France
[2]CNRS, Institut d'Astrophysique Spatiale, Université Paris-Sud 11
Bâtiments 120-121, 91405 Orsay Cedex, France
[3]CNRS, Laboratoire de l'Accélérateur Linéaire, Université Paris-Sud 11
Bâtiment 200, 91898 Orsay Cedex, France
(Dated: June 23, 2009)

We extend the pure pseudo-power-spectrum formalism proposed recently in the context of the Cosmic Microwave Background polarized power spectra estimation by Smith (2006) to incorporate cross-spectra computed for multiple maps of the same sky area. We present an implementation of such a technique, paying particular attention to a calculation of the relevant window functions and mixing (mode-coupling) matrices. We discuss the relevance and treatment of the residual E/B leakage for a number of considered sky apodizations as well as compromises and assumptions involved in an optimization of the resulting power spectrum uncertainty. In particular, we investigate the importance of a pixelization scheme, patch geometry, and sky signal priors used in apodization optimization procedures. In addition, we also present results derived for more realistic sky scans as motivated by the proposed balloon borne experiment EBEX. We conclude that the presented formalism thanks to its speed and efficiency can provide an interesting alternative to the CMB polarized power spectra estimators based on the optimal methods at least on angular scales smaller than ~ 10 degrees. In this regime, we find that it is capable of suppressing the total variance of the estimated B-mode spectrum to within a factor of ~ 2 of the variance due to only the sampling and noise uncertainty of the B-modes alone, as derived from the Fisher matrix approach.

I. INTRODUCTION

The reliable characterization and scientific exploitation of the polarized Cosmic Microwave Background (CMB) Anisotropy signal is one of the main challenges facing the CMB research at the present. The challenge is aimed at by an entire slew of the experimental efforts which are currently on-going [1], being deployed [2, 3, 4, 5, 6, 7] or planned [8, 9]. The majority, if not all, of these experiments will effectively observe only part of the sky. This is either due to observational/hardware limitations, particularly for ground-based and balloon-borne experiments, or the presence of foregrounds, or both. The resulting so-called cut-sky effects will therefore have to be taken into account in the data analysis of most of these experiments.

The focus of this paper is on the CMB power spectrum estimation via the pseudo-spectrum technique [10]. The power spectrum is the most fundamental statistics considered in the context of the CMB fluctuations, which are thought of as being Gaussian or nearly Gaussian distributed. The pseudo spectrum approach provides a computationally quick and flexible framework for estimating the power spectra, which has been extensively and successfully employed in past analyses of the CMB data sets (e.g., [11, 12]). However, it has been long recognized [13] that a straightforward application of the pseudo spectrum technique to cut-sky polarized CMB maps leads to the so-called "E-to-B" leakage, or power aliasing, as a consequence of which the cosmologically important information contained in the CMB B-modes is overwhelmed by the statistical uncertainty of the (much larger) E-modes. To date two techniques alleviating the problem have been proposed: one correcting the leakage on the correlation function level [14] and the other attempting to do so directly on the maps [15]. The latter approach relies on a suitably chosen sky apodization to remove from the map harmonic modes which are neither solely E nor B. In a recent paper [16] it has been suggested that the apodization can be appropriately optimized to bring down the resulting B-mode spectrum uncertainty in-line with the optimal methods, building up a strong case in favor of this approach. In this paper we extend the original formalism of [15] to incorporate it within the framework of the cross-spectra computations and discuss a number of implementation issues, related to derivative computation and apodization choices in actual applications. Later, in this context we discuss the performance of the approach.

We note here that other approaches to the polarized power spectrum, which do not suffer from the leakage problem, exist but are however hampered either by numerical efficiency [17] or convergence [18] issues in particular for the high resolution observations (but see [19]

*Electronic address: julien.grain@ias.u-psud.fr
†Electronic address: tristram@lal.in2p3.fr
‡Electronic address: radek@apc.univ-paris7.fr

for a potential resolution of this problem). Consequently, the pure pseudo-spectrum approach stands out as an important alternative potentially capable of addressing successfully the needs of the future CMB B-mode polarization experiments.

The outline of this paper is as follows. In Sect. II we start from presenting the standard pseudo-power-spectrum technique and the E/B-leakage problem. We also summarize the original approach of Smith (2006) [15] and describe its specific variant as proposed in this paper and which invokes harmonic domain derivatives and the cross-spectrum idea. We conclude that Section with an overview of the numerical implementation of the method. In Sect. III we elaborate on the issues related to the proper sky apodizations. In Sect. IV we consider the performance of the formalism focusing on the role of the apodization and its impact on the quality of the final results, as expressed by the level of the remaining residual E/B leakage and the uncertainty of the estimated power spectra. We summarize the main results of this work in Sect. V.

For definiteness, hereafter, we will reserve the term mask to a binary, pixelized map of a sky assuming only two values 1 and 0, corresponding to an observed or unobserved (or just not to be taken into the analysis) sky pixel respectively. The pixel weights which can assume any nonzero value will be referred to as either apodizations or windows. In this language a mask just defines an observed sky. We will also refer to aliasing involving different spin harmonics as leakage, reserving the term aliasing solely to mixing between different modes of the same type of the same basis functions. Though not strictly precise such differentiations will be useful in the following discussion.

II. NOTATION AND FORMALISM

A. Harmonic decomposition

The CMB (linear) polarization field is completely described by two Stokes parameters, Q and U. Those can be combined into two spin-2 and spin-(-2) fields defined as,

$$P_{\pm 2} \equiv Q \pm iU. \tag{1}$$

For the full sky, the spin fields can be expressed in the harmonic space making use of the spin-weighted spherical harmonic basis

$$P_{\pm 2} = \sum_{\ell m} {}_{\pm 2}a_{\ell m \, \pm 2}Y_{\ell m}. \tag{2}$$

Alternatively, the polarization field can be decomposed into a gradient, E, and a curl, B, part, both of which are coordinate independent. In the harmonic space, these two approaches are related via the following relations,

$$a_{\ell m}^E = -\frac{1}{2}\left({}_{2}a_{\ell m} + {}_{-2}a_{\ell m}\right), \tag{3}$$

$$a_{\ell m}^B = \frac{i}{2}\left({}_{2}a_{\ell m} - {}_{-2}a_{\ell m}\right). \tag{4}$$

From a point of view of physics, the E/B decomposition of CMB polarization may seem more natural as it is directly linked to the primordial cosmological perturbations leading to the presently observed CMB anisotropies. In particular, only primordial gravity waves, and not density perturbations, can create B-type polarization [20]. As a consequence constraining the amplitude of such modes could allow us to study the character of the primordial cosmological perturbations. This simple picture is spoilt somewhat by a presence of an additional B-mode type contribution generated by the gravitational lensing of the E-mode polarization [21] and which needs to be accounted for if the primordial component is to be recovered.

A computation of polarized angular power spectra makes use of spin-weighted spherical harmonics. Those functions, a natural harmonic basis for spin-s fields on the sphere, are derived from the standard (or spin-0) spherical harmonics by applying the spin-raising ($\eth$) or spin-lowering ($\bar{\eth}$) operators [20],

$$
\begin{aligned}
{}_{s}Y_{\ell m} &= \sqrt{\frac{(\ell-s)!}{(\ell+s)!}}\,\bar{\eth}^s Y_{\ell m}, \\
{}_{-s}Y_{\ell m} &= (-1)^s \sqrt{\frac{(\ell-s)!}{(\ell+s)!}}\,\bar{\eth}^s Y_{\ell m},
\end{aligned}
\tag{5}
$$

where hereafter $s \geq 0$. In the following, we will always assume that s is nonnegative. The main properties of the spin-weighted spherical harmonics under different transformation can be found elsewhere, e.g., [20, 22, 23].

Any complex, spin-s field, with ${}_{s}\phi^\dagger = {}_{-s}\phi$ is completely characterized by its projection on the spin-s spherical harmonics, defining its harmonic representation,

$$\pm_s a_{\ell m} = \int d^2\Omega \, {}_{\pm s}\phi \, {}_{\pm s}Y_{\ell m}^\dagger. \tag{6}$$

Just as is the case for spin-2 polarization fields a spin-s field can be also fully described by its E/B harmonic decomposition. Adopting a vectorial notation,

$$\mathbf{\Phi} = \begin{pmatrix} \mathrm{Re}\left[{}_{s}\phi\right] \\ \mathrm{Im}\left[{}_{s}\phi\right] \end{pmatrix}, \tag{7}$$

we can define the E and B spherical harmonic basis as,

$$
\begin{aligned}
{}_{E}\mathbf{Y}_{s,\ell m} &\equiv \mathbf{D}_s^E \, Y_{\ell m} \\
&= \frac{1}{2}\sqrt{\frac{(\ell-s)!}{(\ell+s)!}} \begin{pmatrix} \bar{\eth}^s + (-1)^s\bar{\bar{\eth}}^s \\ -i(\bar{\eth}^s - (-1)^s\bar{\bar{\eth}}^s) \end{pmatrix} Y_{\ell m} \\
&= \frac{1}{2} \begin{pmatrix} {}_{s}Y_{\ell m} + (-1)^s{}_{-s}Y_{\ell m} \\ -i({}_{s}Y_{\ell m} - (-1)^s{}_{-s}Y_{\ell m}) \end{pmatrix},
\end{aligned}
\tag{8}
$$

and

$$
\begin{aligned}
{}_B\mathbf{Y}_{s,\ell m} &\equiv \mathbf{D}_s^B \, Y_{\ell m} \qquad\qquad\qquad\qquad (9)\\
&= \frac{1}{2}\sqrt{\frac{(\ell-s)!}{(\ell+s)!}} \left(\begin{array}{c} i(\eth^s - (-1)^s\bar{\eth}^s) \\ \eth^s + (-1)^s\bar{\eth}^s \end{array} \right) Y_{\ell m}\\
&= \frac{1}{2}\left(\begin{array}{c} i({}_sY_{\ell m} - (-1)^s{}_{-s}Y_{\ell m}) \\ {}_sY_{\ell m} + (-1)^s{}_{-s}Y_{\ell m} \end{array} \right).
\end{aligned}
$$

Here we have defined two differential operators $\mathbf{D}_s^{E(B)}$ which generalize to arbitrary spin the operators used in [13]. The harmonic representation of the field $\mathbf{\Phi}$ in the E/B subspace then reads,

$$
{}_Xa_{s,\ell m} = \int d^2\Omega \, W\left(\Omega\right) \, \mathbf{\Phi} \cdot {}_X\mathbf{Y}_{s,\ell m}^\dagger, \qquad (10)
$$

where $X = E$ or B and we have introduced an arbitrary (scalar) weight, W, which we will later need to account either for the noise present in the data, $\mathbf{\Phi}$, or the cut-sky effects. The dot denotes a standard componentwise dot product. In the special case of CMB polarization, the vector $\mathbf{\Phi}$ is simply the polarization vector defined as,

$$
\mathbf{P} = \left(\begin{array}{c} Q \\ U \end{array} \right).
$$

It can be easily verified that introducing the polarization vector, $\mathbf{P}$ into Eq. (10) and setting $s = 2$ and $W = 1$ over the full sphere, leads to the standard definition of the E and B multipoles.

B. Pseudo power spectra

The pseudo-C_ℓ estimator, *e.g.* [11, 12, 26], in a given bandpower α with a width $\Delta\ell$, is then based on the multipole decomposition Eq. (10) and defined as,

$$
\tilde{\mathcal{C}}_{(i,j)}^{X,\,\alpha} = \sum_{\ell\in\alpha} \frac{\ell(\ell+1)}{2\pi\Delta\ell} \sum_m \frac{{}_X\tilde{a}_{2,\ell m}^{(i)} \, {}_X\tilde{a}_{2,\ell m}^{(j)\,\dagger}}{2\ell+1}, \qquad (11)
$$

where X stands for E or B and i and j distinguish two sets of measurements (maps) of the sky for which the harmonic coefficients have been computed. Throughout this paper, we use $\ell(\ell+1)/2\pi$ as ℓ-weighting while computing the bandpowers. Though, this may not be the most natural choice from the B mode power spectrum perspective, the results presented hereafter are expected to be largely independent on it, given the relatively narrow ℓ–bins used in this work and the fact that the B-mode power spectrum is nearly structure free. For definiteness we will be assuming hereafter that the two different sets of data always refer to the same sky area, though we will

allow for different (nonzero) weights assigned to each of them. The pseudo-spectrum formalism can be used to estimate EE and BB as well as EB spectra. Hereafter, we will use a calligraphic typeface, $\mathcal{C}$, for the estimated and a roman one, C, for the true underlying CMB power spectrum.

Because of the limited sky coverage, and nonuniform, pixel-dependent weights, the above estimator is biased and its average over CMB realizations, $\left\langle \tilde{\mathcal{C}}_{(i,j)}^{X,\,\alpha} \right\rangle$, involves a mixing between different ℓ modes (or bins) and polarization states. The latter can be described by a so-called *mixing kernel* which we will denote as $M_{\alpha\alpha'}^{XX'}$. The unbiased estimator $\mathcal{C}_{(i,j)}^{X,\,\alpha}$ is thus obtained by inverting the following linear system,

$$
\left(\begin{array}{cc} M_{\alpha\alpha'}^{diag} & M_{\alpha\alpha'}^{off} \\ M_{\alpha\alpha'}^{off} & M_{\alpha\alpha'}^{diag} \end{array} \right) \left(\begin{array}{c} \mathcal{C}_{(i,j)}^{E,\,\alpha'} \\ \mathcal{C}_{(i,j)}^{B,\,\alpha'} \end{array} \right) = \left(\begin{array}{c} \tilde{\mathcal{C}}_{(i,j)}^{E,\,\alpha} - N_{(i,j)}^{E,\,\alpha} \\ \tilde{\mathcal{C}}_{(i,j)}^{B,\,\alpha} - N_{(i,j)}^{B,\,\alpha} \end{array} \right).
$$
(12)

Here $N_{(i,j)}^{X',\,\alpha}$ is the noise contribution to the estimated pseudo-spectra and vanishes, whenever the noise in the two data sets is not correlated, as for example, in a case of two data sets produced by two different experiments or two uncorelated detectors of a single experiment. In fact, we will always assume that $N_{(i,j)}^{X',\,\alpha'} = 0$, whenever $i \neq j$, and refer to the latter cases as pseudo cross-spectra. We will call a pseudo-spectrum computed from a single data set, i.e., $i = j$, a pseudo autospectrum. We note that in this case the noise term is usually nonvanishing and has to be estimated from the data and/or our knowledge of the experiment to allow the estimator in Eq. (11) to be correctly unbiased. This emphasizes one of the biggest advantages of the cross-spectrum based estimators, which do not need such a correction. However, the mixing kernel, $M_{\alpha\alpha'}^{XX'}$ will depend on the window functions applied to each of the data sets and therefore be somewhat more cumbersome to calculate in the cross-spectrum case. Moreover, as we will discuss that later on, some of the most appealing choices for the window function will often require the knowledge of the noise in the data. Nevertheless, as the latter will be used only for the power spectrum error estimation and not for the computations of the spectral estimates, lower precision in modeling the noise may be sufficient. In the very least, in a case of real data the cross-spectra provide a useful and handy diagnostic in the analysis of any data set [27], while for a discussion as the one presented here the cross-spectra allow to simplify the analysis by avoiding any explicit treatment of the noise bias, while retaining all the other aspects essentially unchanged.

The mixing matrix can be computed as,

$$M_{\alpha\alpha'}^{XX'} = \sum_{\ell\in\alpha}\sum_{\ell'\in\alpha'}\frac{\ell(\ell+1)}{\ell'(\ell'+1)\Delta\ell}\sum_m\frac{1}{2\ell+1}\sum_{ij}\tilde{\mathbf{Y}}_{X,\ell m}^{\dagger}(x_i)\mathbf{S}_{\ell'}^{XX'}(x_i-x_j)\tilde{\mathbf{Y}}_{X,\ell m}(x_j), \tag{13}$$

where x_i denotes a position of an i-th pixel. The above expression is fully general, as $\tilde{\mathbf{Y}}_{X,\ell m}$ can be any function used to project the Stokes parameters maps into the harmonic E and B subspaces. In particular, for the standard spectra the windowing of the data with a window W can be taken into account by setting,

$$\tilde{\mathbf{Y}}_{X,\ell m} \equiv W\,\mathbf{D}_2^X\,Y_{\ell m}. \tag{14}$$

The specific expressions for $M_{\alpha\alpha'}^{XX'}$ relevant for the case of standard pseudo spectra can be found elsewhere [12].

The $\mathbf{S}_\ell^{XX'}$ matrix is a convolution kernel relating the pixel-domain correlation functions of the Q and U Stokes parameters to the E and B angular power spectrum (the explicit link between two-point correlation functions and power spectra for polarization fields can be found in the Appendix of [17]). We point out that in some applications considered in the following the window, W, and therefore also the functions $\tilde{\mathbf{Y}}_{X,\ell m}$ will all depend on the bin number, α, (see Eq. (12)).

In general, the kernel, $M_{\alpha\alpha'}^{X,X'}$, will mix different ℓ-modes as well as E and B polarization types. The consequences of this fact and the conditions, which ensure the diagonality of the kernel in the polarization states are discussed in the following Sections.

C. The E/B leakage problem

The operators defined in Sect. II A, valid for any spin-s complex field, define two orthogonal subspaces as,

$$\mathbf{D}_s^{E\dagger}\cdot\mathbf{D}_s^B = 0.$$

However, the E/B subspaces defined by $_{E/B}\mathbf{Y}_{s\ell m}$ are orthogonal (in the sense of Eq. (10)) only on the full sky. This means that if the polarization field is known on a limited part of the celestial sphere, the reconstruction of E and B modes by projecting the measured polarization on the $_{E/B}\mathbf{Y}_{2\ell m}$ basis is "nonpure", i.e., the estimated E or B multipoles receive a contribution from both E and B modes. This applies in particular to the pseudo-power spectra, $\tilde{\mathcal{C}}^{EE}$ and $\tilde{\mathcal{C}}^{BB}$, as defined in Eq. (11), each of which will include a contribution from the other mode. Though such a mixing can be removed on average, the leaked modes will still contribute to the variance of the estimated power spectra. Therefore, if the power contained in one of the polarization states is much higher than in the other, as is the case of, for example, the CMB anisotropies, the error bars estimated for the mode with lower power will be drastically exaggerated due to sample variance of the leaked contribution.

The same conclusion can be arrived at by starting from Eq. (10) which can be viewed as a harmonic decomposition of the $(W\mathbf{P})$ field. Multiplication by the window function, W, is a convolution not only in the harmonic space but also in the E/B subspace, turning thus $\tilde{a}_{\ell m}^B$ into a mixture of the true B and E CMB multipoles.

This last observation highlights the fact that the standard pseudo-C_ℓ technique may be in practice unable to provide good estimates of the B-mode angular power spectra in particular from data sets coming from small-scale CMB experiments. Moreover, even for satellite experiments for which angular power spectra can be estimated from a significant fraction of the sky, the discussed effects may be significant enough not to be negligible.

We can look at the E/B leakage problem also from the mixing matrix perspective. The lack of orthogonality of E and B type of spherical harmonics results in the coupling between the E and B pseudo power spectra. For this reason the inversion in Eq. (12) needs to be performed simultaneously for both E and B spectra to allow for unscrambling the modes with the help of non-vanishing off-diagonal EB and BE blocks of the mixing matrix. The resulting expression for the power spectrum estimate, derived in the case of the B-mode spectrum reads now,

$$\begin{aligned}
\mathcal{C}_{(i,j)}^{B,\,\alpha} &= \left(M_{\alpha\alpha'}^{diag} - M_{\alpha\alpha'}^{off} M_{\alpha\alpha'}^{diag\,-1} M_{\alpha\alpha'}^{off}\right)^{-1}\\
&\times \left[\left(\tilde{\mathcal{C}}_{(i,j)}^{B,\,\alpha'} - N_{(i,j)}^{B,\,\alpha'}\right)\right.\\
&\left.- M_{\alpha\alpha'}^{off} M_{\alpha\alpha'}^{diag\,-1}\left(\tilde{\mathcal{C}}_{(i,j)}^{E,\,\alpha'} - N_{(i,j)}^{E,\,\alpha'}\right)\right].
\end{aligned} \tag{15}$$

In the ensemble average sense the above expression is unbiased as a result of a subtle cancellation of the E mode power present in the pseudo-B and E spectra. Such a cancellation does not however apply to the variance of the estimator and as a result the variance of the spectra of one type will include a contribution from the other.

In the following we will quantify the E/B leakage using both these perspectives, i.e., either looking at an excess variance of the estimated B-mode power spectrum or at the magnitude of nonvanishing, off-diagonal blocks of the mixing matrix.

D. Pure pseudo-C_ℓ estimators for cross-spectrum

1. Definition

Pseudo-C_ℓ estimators of the polarization power spectra which do not mix E and B modes can be constructed in

projecting the polarization fields on the "pure" E and B subspaces [13]. Pure E and B multipoles on a partial sky are defined as follows (the vector $\mathbf{P}$ now represents the polarization signal as measured on the sky) [15]:

$$_E\tilde{a}_{\ell m} = \int dx \left[\mathbf{D}_2^E \left(W(x) Y_{\ell m}(x) \right) \right]^\dagger \cdot \mathbf{P}(x), \quad (16)$$

$$_B\tilde{a}_{\ell m}^B = \int dx \left[\mathbf{D}_2^B \left(W(x) Y_{\ell m}(x) \right) \right]^\dagger \cdot \mathbf{P}(x). \quad (17)$$

with W a spin-0 window function satisfying the Dirichlet and Neumann conditions on the boundary of the observed sky region. Such conditions on the window function are mandatory for the estimated multipoles to be free of a E/B leakage due to partial sky. This can be easily shown in the pixel-domain. On applying an integration by parts twice to Eqs. (16) & (17) and taking into account the boundary properties of W, we can rewrite the estimated multipoles as [15, 16],

$$_E\tilde{a}_{\ell m} = \int dx \, W(x) Y_{\ell m}(x) \, \mathbf{D}_2^{E\dagger} \cdot \mathbf{P}(x), \quad (18)$$

$$_B\tilde{a}_{\ell m} = \int dx \, W(x) Y_{\ell m}(x) \, \mathbf{D}_2^{B\dagger} \cdot \mathbf{P}(x). \quad (19)$$

Given that the operator $\mathbf{D}_2^E$ ($\mathbf{D}_2^B$) filters out all the B (E) modes, the multipole decomposition defined above indeed should be free of any E/B mixing. (We point out that the above harmonic decomposition corresponds to the decomposition on a partial sky of the so-called χ fields defined in [16].) A simple way to understand this point is to notice that applying $\mathbf{D}_2^B$ to the polarization field corresponds to a local filtering of the E modes whereas projecting on the $_B\mathbf{Y}_{2,\ell m}^\dagger$ basis corresponds to a global filtering. Because of this difference, the latter of these two techniques suffers from partial sky effects whereas the former one does not. In other words, if the window, W, satisfies the Dirichlet and Neumann boundary conditions, the multipoles estimated in Eqs. (16) and (17) are pure and the mixing kernel, $M_{\alpha\alpha'}^{XX'}$, is a block diagonal matrix in the E/B subspace, i.e.,

$$M_{\alpha\alpha'}^{XX'} = M_{\alpha\alpha'}^{XX} \delta_{X,X'}. \quad (20)$$

This last condition, intuitively obvious given our previous discussion, can be formally derived from Eq. (13) defining,

$$\tilde{\mathbf{Y}}_{X,\ell m} \equiv \mathbf{D}_2^X W Y_{\ell m}. \quad (21)$$

A more detailed discussion of the mixing matrix and related issues is presented in the next Section and Appendix A.

2. Numerical implementation

a. *Pure multipoles.* In the actual implementation, we do not make a direct use of Eqs. (18) & (19) as this would require taking numerical derivatives of noisy maps. We rather calculate the pure multipoles using Eqs. (16) & (17). We refer hereafter to this way of performing calculations as the spin-weighted approach. Consequently, $\mathbf{D}_2^{E/B}$ operators need to be applied directly only to products of the spherical harmonics times the window function. This is an approach also adopted in [15, 16]. Our numerical implementation proceeds in two steps. First, we define two spin-weighted windows,

$$W_1 = \eth W \quad \text{and} \quad W_2 = \eth^2 W, \quad (22)$$

and construct three new fields,

$$\begin{aligned} \tilde{P}_2 &= W(Q + iU), \\ \tilde{P}_1 &= W_1^\dagger (Q + iU), \\ \tilde{P}_0 &= W_2^\dagger (Q + iU). \end{aligned} \quad (23)$$

$\tilde{P}_s$ is a spin-s field, since $W_s^\dagger = W_{-s}$ has spin $(-s)$ and $(Q + iU)$ – spin 2. Then, on the second step, we calculate the pure multipoles as linear combinations of the E and B multipoles of the three new fields. The latter can be derived from Eqs. (16) & (17) with the help of Eqs. (8) & (9). For instance, in the case of the B-modes, the pure estimated multipoles read,

$$_B\tilde{a}_{\ell m} = B_{2,\ell m} + 2\sqrt{\frac{(\ell-2)!(\ell+1)!}{(\ell+2)!(\ell-1)!}} B_{1,\ell m} + \sqrt{\frac{(\ell-2)!}{(\ell+2)!}} B_{0,\ell m}, \quad (24)$$

with $B_{s,\ell m}$ denoting the B-type multipoles of the $\tilde{P}_s$ field as in Eq. (10). Numerically the entire computation comes down to an efficient calculation of the spin spherical harmonic transforms. We comment on those in more detail later in this Section. Equation (24) offers another insight into the pure estimator formalism. The first term on the *rhs* of this equation corresponds to the standard estimator, while the two following ones act as counter-terms removing the leaked E-mode. Indeed we have found that they are anticorrelated with the first term as expected in such an interpretation.

b. *Numerical derivatives.* The knowledge of the three spin window functions is needed for numerically computing the pure multipoles. As will be clear in the next Section, there are cases where only the spin-0 window is known. Therefore, we need an efficient way to derive the spin-1 and spin-2 windows from the spin-0 one, which in turn requires a numerical calculation of window derivatives. Hereafter we will perform this computation in the harmonic space,

$$W_s = \eth^s W \to w_{s,\ell m} = \sqrt{\frac{(\ell+s)!}{(\ell-s)!}} w_{\ell m}. \quad (25)$$

The final spin-1 and spin-2 windows are then obtained by transforming back to the pixel-domain. This technique allows us to use simple analytic window functions on complicated sky coverage, for which only the spin-0

part can be easily computed. It also ensures that the conditions given in Eq. (22) are fulfilled. By doing this computation in harmonic space, we may introduce some ringing due to the pixelization (or gridding) of the sphere. Indeed, we have found some minor artefacts of this sort when comparing numerically computed windows with the respective analytic expressions in cases when the latter could be calculated, e.g., in cases with apodizations as in Eqs. (30) & (31) and for simple observed sky patches. It appears, however, that such ringing does not seem to affect our ability to control the E/B leakage at scales where this control is mandatory. For instance, for a 20×20 degree square patch pixelized with HEALPix pixels with $N_{side} = 512$, the residual, leaked E-modes calculated using the numerical spin windows reproduce closely those obtained analytically all the way up to $\ell \sim 400$. For higher values of ℓ, the use of numerical computation leads to a slight increase of the residual leakage. We expect that to be always the case, independently of the apodization length as long as it is sufficiently greater than the pixel size.

In general the discrepancy between the exact and numerical spin windows depends on patch geometry. For example, the discrepancy is reduced for a spherical cap as compared to a square patch. Its magnitude will also in general depend on the assumed pixelization and in particular may be greatly enhanced if the considered region is in a particular position and/or orientation with respect to the pixelization scheme. However, whenever necessary, such a discrepancy can be reduced or at least shifted to higher ℓ-values by going to a higher resolution. We demonstrate some of these effects in Sect. IV.

c. Mixing kernel. The mixing kernel can be expressed as a sum over the ℓ-modes in the harmonic space of products of the harmonic coefficients of the spin window functions and $3j$ Wigner coefficients [11]. This is done by using Eq. (21) in Eq. (13) and subsequently expanding the window function in harmonic space. We present relevant formulae in Appendix A (see also [15] for an alternative derivation of those kernels). They generalize the standard expressions for the ordinary pseudo spectra [24, 25]. In our numerical computations we implement these formulae and perform all the required calculations directly in the harmonic domain. This is in contrast to the previous study of [15], which strives to make computations predominantly in the pixel-domain.

As we mentioned before for any scalar window function we can calculate the harmonic coefficients of the spin windows by using the relation in Eq. (25). If we then use these in the computation of the mixing kernel, Eqs. (A4) & (A5), we find that the off-diagonal blocks of the mixing matrices vanish within the numerical precision. This simply reflects the fact that the pure estimators are indeed "pure" and do not mix E and B modes. This is however also not realistic. In practice, we expect that some residual E/B leakage will be present due to the sky discretization – an effect not corrected by the pure pseudo spectra formalism and not characterized properly
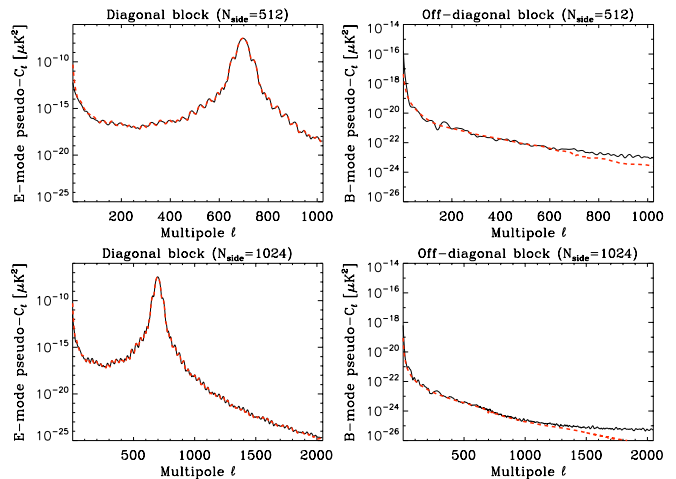

FIG. 1: Dashed (red) curves show a single column of the mixing kernel at $\ell_0 = 700$ computed using our numerical approach as described in the text, for the diagonal, left, and off-diagonal blocks, right, and two different resolutions: $n_{side} = 512$, top row, and $n_{side} = 1024$, bottom. The overlapping solid (black) lines depict the EE, left, and BB, right, pseudo-spectra computed from Q and U maps containing only E-mode power at $\ell_0 = 700$ and apodized with analytic window functions satisfying the boundary condition and constrains in Eq. (22). The apodization length is set to 7 degrees. A very good agreement seen in all the cases validates the proposed numerical approach. The off-diagonal elements in the left panels are due to the cut-sky mode mixing, while the nonzero values in the right panels are due to a residual pixel-induced E/B leakage.

in the above mentioned calculation of the mixing matrix. To account for the pixelization effects we propose to compute the mixing kernel using coefficients of each of the spin windows calculated via numerical spherical harmonic transforms (SHT) of the pixel-domain representation, Eqs. (A7) & (A8). We do so even in the cases, when the window derivatives are computed numerically and thus are directly available in the harmonic domain. In such cases, our procedure corresponds simply to taking first an inverse SHT of the spin windows as given in the harmonic domain. This produces their pixel-domain representation, which we transform back to harmonic domain via a direct SHT. Because of sampling effects such an operation does not reproduce precisely the input coefficients. The new harmonic coefficients of the windows, when used now to calculate the mixing matrix, lead the higher values of the off-diagonal block elements, which, though still small when compared with the diagonal elements of the diagonal blocks, are not any more numerically zero. This can be understood from Eq. (24): because of gridding effects during the SHT, the multipoles involved in the pure estimators are not computed exactly and the precise cancellation of the leaked E-mode is not perfect.

As those are the pixel-domain representation of the windows, which are used in the pure pseudo-spectra computation, we may expect that the observed change by

explicitly taking the SHT of each spin windows while computing the mixing kernel, goes in the right direction, properly reflecting the correlation pattern between the modes of the pure spectra. We can test this expectation with the help of numerical experiments. We first note that we can compute any of the columns of the mixing matrix by computing the pure E and B pseudo-spectra of the sky signal containing power only in one selected harmonic mode $\ell = \ell_0$. Examples of such computations are shown in Fig 1. There we have allowed for nonzero power only in a single $\ell_0 = 700$ mode of the E polarization. The derived pure E and B pseudo spectra are contrasted with the corresponding columns of the EE and BE blocks of $M_{\ell\ell'}$, both computed using the procedure outlined before.

For a square patch assumed in the calculation visualized in Fig. 1, the overall agreement between the pseudo spectra and the column of the kernel is indeed very good. For the off-diagonal block some differences can be seen only at the high-ℓ end with the mixing kernel underestimating the residual leakage. The discrepancy appears mostly benign as it affects the multipoles for which the leaked signal is found to be negligible.

We thus conclude that the simple way of incorporating the pixel effects as proposed here works indeed very well everywhere, where the leakage is significant and thus requires a sufficiently accurate treatment. This also demonstrates that the method of computing the mixing matrix advocated here adequately describes the actual correlations between the ℓ-modes of the computed polarized pseudo-spectra. For more relevant examples of such a comparison see Sect. III A and Fig. 2.

We note that in general different window functions can be used for the estimation of the power in different ℓ-bins. Indeed, as we discuss that later on, it is a common occurrence in practical applications in which we usually strive for the smallest attainable uncertainty on a final power spectrum. In such cases, we calculate the kernel row-by-row, i.e. bin-by-bin, using an appropriate window assigned to each of them. The kernel computed in this way is later used to solve the inverse problem in Eq. (12).

d. Spherical harmonic transforms. At the core of the numerical implementation of the formalism presented in the paper is a calculation of the spin spherical harmonic transforms. For this purpose we utilize here a publicly available, Scalable Spherical Harmonic Transforms (S^{2}HAT) library [30]. This parallel library allows for an efficient computation of arbitrary spin transforms for a general class of sky griddings and pixelization in which the pixels (grid points) are distributed along the constant declination lines. These include therefore the most frequently used in the CMB work, e.g., HEALPix [31], GLESP [32], ECP [33] as well as some others. The spin spherical harmonics are computed using the recurrence relations as listed in [34], which are appropriately rescaled to avoid numerical under- and overflows. The library is designed to treat optimally the cases with only a partial sky coverage and is therefore particularly suitable

for our purposes here. It is parallelized using the Message Passing Interface (MPI) and distributes over the processors both the pixel (maps) and harmonic (harmonic coefficients) domain objects permitting to achieve high resolutions with modest memory resources available per single processor.

III. SKY APODIZATIONS

Performance and in a result applicability of the pure (cross) spectra formalism is clearly dependent on our ability to compute a suitable apodization for any specific CMB polarization experiment. In this Section we introduce a number of proposal functions, which could be used as apodizations in calculations of the pure (cross) spectra. The discussed functions range from ones derived as a result of optimization procedures and therefore often numerically involved, to ones proposed somewhat *ad hoc* but quick and easy to calculate. We will discuss their relative merit only briefly here, postponing a more thorough comparison of their performance to the following sections.

A. Variance-optimized apodizations

In realistic applications we would like to have an apodization function which would ensure that the total (i.e., noise plus sample variance) uncertainty of the recovered power spectra is (nearly) minimal. In the case of the auto pure pseudo spectra a general ansatz for such optimized weighting has been proposed recently by Smith and Zaldarriaga [16]. It is derived as a condition on the apodization window required to minimize the difference between the auto pseudo-spectrum of a map and its optimal counterpart. Those authors have obtained a following relation for the (nearly) optimal weights,

$$\sum_{i=1}^{N_{obs}} \mathbf{C}_{ij} \mathbf{P}_{ij}^{(\alpha)} W_i = 1, \qquad \text{for all } j. \qquad (26)$$

Here the summation is made over the N_{obs} observed pixels. The matrix $\mathbf{C}$ is the covariance matrix of the data and $\mathbf{P}$ is a geometrical matrix which comes from the projection on harmonic space (for example the Legendre polynomials matrix in the case of temperature). The optimized weight function can be then obtained by a numerical inversion of the above linear system for each bin α. It can be shown [16] that such optimized windows converge towards inverse square-noise weighting in the noise-dominated case. This is a regime usually reached at the high-ℓ end of the estimated power spectrum. We refer the reader for more details about the derivation of this relation to [16].

For the computation of the cross-spectrum we need to compute an apodization for each of the available maps.

The approach of [16] has to be therefore extended to accommodate such cases. This can be done in a number of ways. For instance, using Eq. (26) we can compute an apodization for each map considering them separately or together. Alternately, we could apply a reasoning similar to that of [16] but in the context of the cross-spectrum to derive a relation analogous to Eq. (26) but applicable to the cross-spectra. Though this last approach may seem the most appropriate for the case at hand it unfortunately results in an expression, Eq. (B14), which is computationally rather involved. The two first options mentioned above are numerically more tractable, though they produce windows, which minimize the difference between the pure pseudo-spectrum estimator and its optimal counter-part on the auto spectrum level. The first of these two approaches uses for that purpose the auto spectrum of the map for which the apodization is sought, while the second one – the single auto spectrum of all the maps considered. Clearly, which one of the two methods is more appropriate will depend on the specific goals of the entire estimation procedure. We note for example that in the case of two maps with the same noise properties the difference between the two optimization approaches is limited only to a different overall factor determining the level of the noise with respect to the sky signal (see Appendix B for more details). Hereafter, we will utilize the first option and solve Eq. (26) as in the pure pseudo auto spectrum case for each of the two maps. We will support this choice in the following Sections with the help of numerical experiments.

Equation (26) can be specialized to a particular case of polarization when the input data used are composed of the Stokes Q and U maps. The geometrical matrix $\mathbf{P}_{ij}^{(\alpha)}$ needs to be then extended into the spin subspaces and nonzero spin windows introduced. Their implicit definition consequently reads,

$$\sum_{i=1}^{N_{obs}} \mathbf{C}_{ij} \sum_{s=0}^{2} \mathbf{P}_{i+N_{obs}s,j+N_{obs}s'}^{(\alpha)} W_{i+N_{obs}s} = \mathbf{e}_{j+N_{obs}s'},$$

(27)

where an additional summation over spins of the windows is added. The geometrical matrix, $\mathbf{P}^{(\alpha)}$, describing the projection on spin-weighted spherical harmonics, is now explicitly dependent on the spin value, s, and in general is not diagonal in the spin subspace. The vector $\mathbf{e}$ is a $5N_{obs}$ vector whose first N_{obs} entries are equal to one, corresponding to the normalization of the spin-0 component, and all the others are set to zero [16].

The approach can be also looked at as an attempt to minimize the total, signal plus noise, power aliased to a given bin either from the other adjacent bins or the other modes (e.g., E/B leakage). This viewpoint is referred to as the variational interpretation in [16]. Consequently, in general there is always a different window function calculated for each bin as a result of this procedure. This has two consequences. First of all, one has to deal with a set of windows optimized for each bin. Second, this means that one has to define each bin parameters before
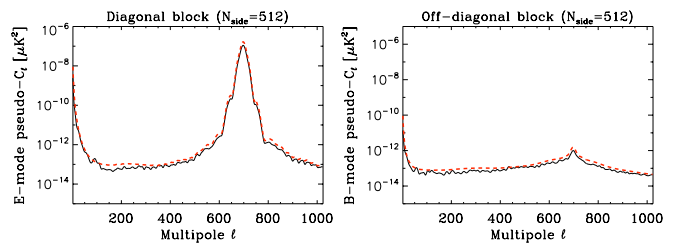

FIG. 2: As figure 1, but computed assuming numerically optimized pixel-domain window functions and a more realistic sky patch as could be observed by CMB B-mode balloon-borne experiments, such EBEX, (see Fig. 21 and Sect. IV for more details).

performing the optimization procedure.

Below we discuss implementation of this kind of optimization procedure in both pixel and harmonic domains.

1. Pixel-domain implementation

The matrix operations in Eq. (27) should be done preferably in a pixel-domain. That allows to consider a full potential complexity of the map noise properties as well as arbitrary masks with complicated boundaries and/or holes due to point-sources extraction, rendering such an implementation very flexible.

In the context of the pure pseudo-spectrum estimators, one should in principle ensure that the spin window functions fulfill the relation in Eq. (22) as well as boundary conditions. However, hereafter following [16], we relax these constraints in the numerical implementation of the procedure. This allows for greater flexibility and therefore enables finding window functions better suited for our needs. Nevertheless, this simplification has a major consequence: the pure pseudo-C_ℓ estimator is not perfectly pure and will unavoidably introduce some E/B leakage in addition to the one induced by the pixelization itself. This residual leakage may become problematic in the low-ℓ region where sampling variance is the major contribution to the error bars. We note however that in the high signal-to-noise ratio regime the resulting windows are expected to nearly fulfill all the constraints, even if those are not explicitly imposed. This is because the total variance minimization approach used in [16] to derive Eq. (26) implies that the solution for which the counter-terms in Eq. (24) nearly cancel the leaked signal included in the first term is preferred over those for which three terms are independent, which unavoidably boosts the variance. In the noise-dominated regime, the leaked E/B signal is quickly becoming subdominant and therefore it will be the noise variance, not the leaked E signal, which will be the major target of the optimization. In this case, the perfect apodization of the window, i.e., as required by the pure formalism, may not be the best solution, which should gradually, with a decreasing signal-to-noise ratio, approach the standard inverse noise

weighting.

We point out here that the window function satisfying Eq. (26) ensures only that the uncertainty is minimized on the pseudo-spectrum not on the power spectrum level. In particular, if the mixing matrix is not block diagonal, i.e., some residual E/B leakage allowed, the final power spectrum uncertainty may not be strictly at its minimum. Nevertheless, though some E/B leakage may be indeed present, it will be correctly accounted for in our calculation of the mixing matrix, and therefore consistently treated in the presented method. We show in Fig. 2 that the computational approach to the calculation of the mixing matrix proposed here, and already discussed in the case of a simple sky patch and analytic window, fares indeed very well also in the case with the nearly optimized window functions.

The fact that Eq. (27) is solved relaxing the derivative constraints between the spin-s and spin-0 components does not reduce it to a set of three independent equations – one for each spin-weighted window. This is because the geometrical matrix $\mathbf{P}_{ij}^{(\alpha)}$ is not diagonal in the spin subspace. Consequently all three windows have to be solved for simultaneously. In order to do so, following [16] we have implemented an iterative Preconditioned Conjugate Gradient (PCG) solver [35]. The numerical cost of such an operation is high and on order of $\mathcal{O}\left(n_{iter}\, n_{pix}^2\right)$, if no symmetries are present. A number of the iterations, n_{iter}, is usually as high as a few hundreds and it is strongly dependent on a dynamic range spanned by the pixel noise amplitudes. Nonetheless, for small-scale experiments with the help of a proper thresholding the computational load is clearly within the limits of available presentday (super)computers. We note that for the high signal-to-noise cases imposing explicitly the constraints on the window functions could reduce the number of iterations required for the procedure to converge. We have not however attempted to include such an optimization in the current version of the software in a general case (see however the homogenous noise discussion in the next Section).

To illustrate the results of our numerical optimization, in Fig. 3 we show three windows computed for a spherical cap with homogeneous noise and two different ℓ-bins. The noise level is set to 5.75 μK-arcmin, the E-mode signal is obtained with the WMAP 5-year cosmological parameters [36] and the B-mode signal includes lensing-induced and primordial B-mode ($T/S = 0.05$). The radius of the cap is roughly 11 degrees, leading to a significant level of E-to-B leakage in the represented bins. The resulting window functions are well but not perfectly apodized as they do not precisely vanish at the edge of the cap. However, as expected, the apodization length, defined as a width of the boundary layer, see Eq. (30), decreases for bins with higher ℓ. This is in agreement with the discussion of [15] concerning the choice of apodization length for homogeneous noise cases.
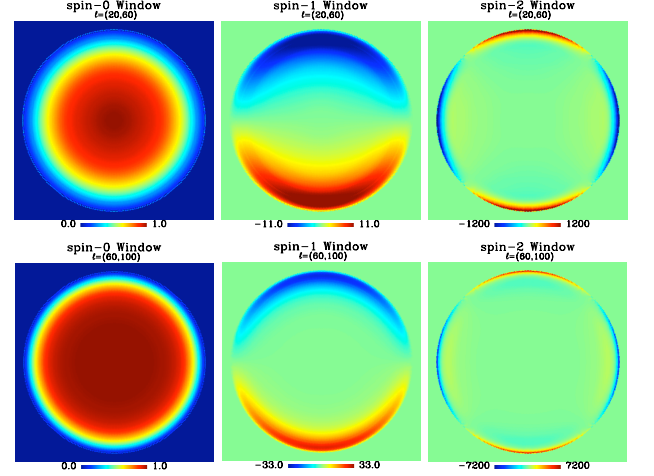

FIG. 3: From left to right: spin-0, spin-1 & spin-2 spin-weighted optimized windows for two bins $\ell \in (20, 60)$ and $\ell \in (60, 100)$ are shown in the upper and the lower panels respectively. In all cases only the real parts of the spin-1 and spin-2 windows are displayed. The noise level is set to 5.75 μK-arcmin, the E-mode signal is the WMAP 5-year best fit [36] and the B-mode signal includes lensing-induced and primordial $B-$mode ($T/S = 0.05$). The sky patch is a cap of a radius of 11 degree leading thus to a significant amount of E-to-B leakage. Note that in the middle (left) panel, the color stretch is 3 (6) times bigger for the case with $\ell \in (60, 100)$ than for $\ell \in (20, 60)$.

2. Harmonic domain implementation

If the noise in the pixel-domain is uncorrelated and homogeneous over the observed sky patch, the overall optimization procedure can be simplified by performing the calculations in the harmonic space. That can be done most straightforwardly starting from Eqs. (18) & (19) rather than Eqs. (16) & (17) and therefore results in windows for which the boundary conditions, Dirichlet + Neumann, are explicitly imposed.

We begin deriving the average B-mode pseudo-spectrum (see Appendix C),

$$\left\langle \tilde{C}_\ell \right\rangle = \sum_{\ell''m''} C_{\ell,\ell''}^\chi w_{0,\ell''m''} w_{0,\ell''m''}^*, \qquad (28)$$

where,

$$\begin{aligned} C_{\ell,\ell''}^\chi &= \sum_{\ell'} \frac{2\ell'+1}{4\pi} \frac{(\ell-2)!(\ell'+2)!}{(\ell+2)!(\ell'-2)!} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2 \\ &\times \left(C_{\ell'}^B B_{\ell'}^2 + \sigma^2 \right). \end{aligned}$$

Here σ^2 denotes the noise power spectrum, B_ℓ – the effect of an azimuthally symmetric beam. The power spectrum $C_{\ell,\ell''}^\chi$ represents the total aliased signal and noise power leaking from the ℓ' bin into the ℓ one.

The optimized apodization is then given by applying the variational principle to $w_{0,\ell''m''}$ with three external constrains, which have to be imposed: the window

has to be normalized within the observed region and together with its first derivative, they have to vanish at the contour of the observed region. The computation then consists of solving a differential equation with the appropriate Dirichlet + Neumann boundary conditions thus leading to the final expression for the optimized window given by,

$$w_{0,\ell''m''} = \frac{1}{C^\chi_{\ell,\ell''}} \left(\lambda \mathcal{M}_{\ell''m''} + \mu \mathcal{P}_{\ell''m''} \right). \qquad (29)$$

Here $\mathcal{M}_{\ell''m''}$ and $\mathcal{P}_{\ell''m''}$ are the harmonic representation of the mask and the contour and λ and μ are two constants given in Appendix C, which both depend on the signal and noise power.

In practice the optimization proceeds in two steps. First, Eq. (29) is used to provide the scalar component of the optimized window. This requires taking the spherical harmonic transform of the mask and of its contour and calculating the weights $C^\chi_{\ell,\ell''}$. Subsequently, the spin-1 and spin-2 windows are derived from the scalar one in the harmonic space using the numerical tools. This procedure requires taking the SHT of the binary mask and the one of the contour of this mask. Because this last SHT is prone to a numerical error (as it corresponds to calculating an integral of a pixelized line), the resulting scalar window function is not perfectly apodized. As a consequence, the window derived in this way does not ensure that the estimator is perfectly (or sufficiently) pure and may therefore reintroduce E/B leakage. We point out that unlike the pixel-domain implementation for which the level of leakage is by design smaller than the noise level, and therefore under control, the extra-leakage introduced is here not controlled and thus may or may not be smaller than the noise. To cure this problem, we add an extra apodization to the scalar window, W_0, before computing its first and second derivatives. This extra-apodization is obtained by multiplying the spin-0 window by the C^2-window with a small apodization length (see the following Section). This makes our window optimal throughout the entire mask (its shape is driven by the optimization procedure) except very close to the boundary, where its shape is driven by the extra apodization.

The results of the numerical implementation for a spherical cap with homogeneous noise are shown in Fig. 4. These windows show the same behavior as the windows obtained from the PCG with respect to bins and signal level. In particular, we find that the apodization length decreases for higher bins. The shapes of spin-1 and spin-2 windows computed in the pixel- and harmonic domains are however significantly different. Because of the extra-apodization and numerical issues, the windows obtained in the harmonic space in a way proposed here are not any more an exact solution of the optimization problem fulfilling the boundary constraints. Nevertheless they provide an easy and quick-to-compute apodization, which one may hope will result in spectral variances not substantially larger than the one produced by the computationally heavy pixel-domain optimized windows de-
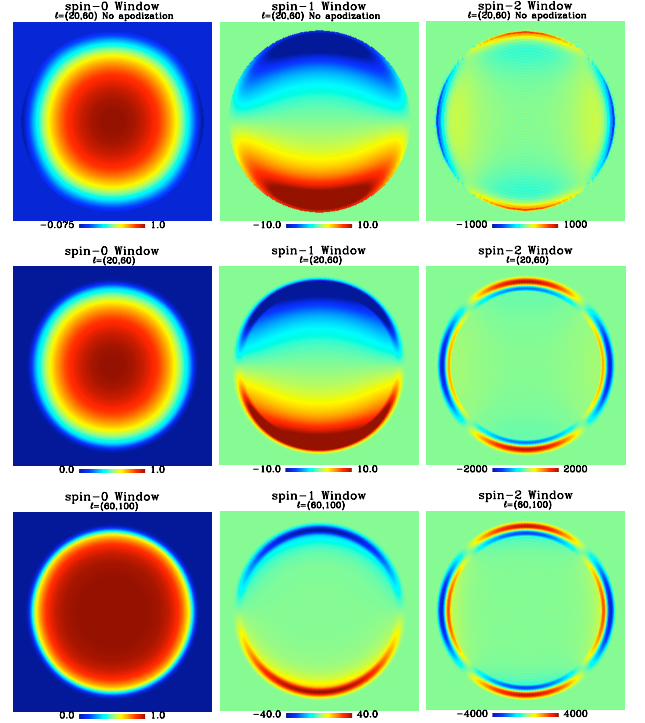

FIG. 4: Left to right: spin-0, spin-1, & spin-2, spin-weighted optimized windows computed in the harmonic space for $\ell \in (20, 60)$ without, *top row*, and with, *middle row*, the extra-apodization as discussed in the text. We show here only the real parts of the spin-1 and spin-2 windows. The color stretch of the right panel is 2 times bigger from top to middle row. *Bottom:* The same as middle row but optimized for a bin $\ell \in (60, 100)$. In the middle and right panels, the color stretch is set to be 4 (2) times bigger, respectively, as compared to the corresponding panels of the middle row. In all the shown cases the B-mode signal includes lensing and primordial contribution for $T/S = 0.05$ and all the other parameters are set as in Fig. 3.

scribed below. We will test these expectations against simulations in Sect. IV.

3. Sky signal prior and the optimized windows

The computations of the properly optimized window functions requires some prior knowledge about the signal expected in the data. In the context of the small-scale CMB observations at the low and intermediate ℓ the optimization is mostly between the variance of the E polarization signal leaked to B and the assumed (and presumed known) noise level. Hence this is the level of the E-mode polarization, which is of primary importance. As polarization experiments measure usually both modes of polarization, therefore a high quality internal constraint on the E-mode power should be available for an experiment aiming at the B-mode detection. At the higher-ℓ end a reasonable guess about the B-mode power is
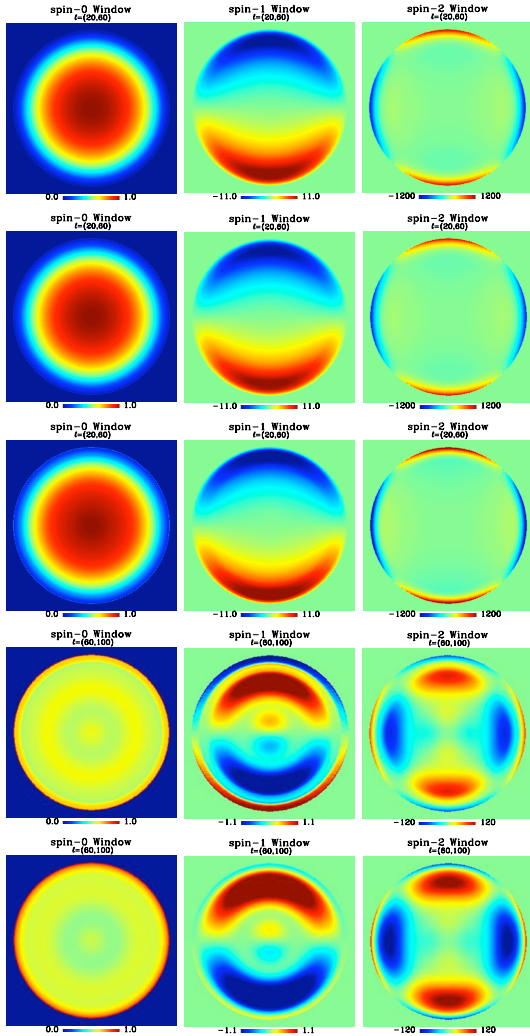
FIG. 5: Spin-0, spin-1, & spin-2 optimized window functions, left to right, computed for the $\ell \in (20, 60)$ bin for three different assumptions about the B-mode signal: no B-mode (first row), lensing-induced B-mode only (second row), lensing-induced plus primordial B-mode for $T/S = 0.05$ (third row). In all these three cases, the E-mode signal is obtained using the cosmological parameters constrained by WMAP 5-year data [36]. In the fourth row the E-mode signal amplitude is set to a hundredth of that value, i.e., on par with the B-mode level, and no E-mode signal is assumed in the fifth row. In these last two cases, the B-mode is the same as in the third row. The noise and patch properties are the same as in Figs. 3 & 4. The shape of the windows depends weakly on the assumed B-mode, but clearly more strongly on the E-mode. For visualization purposes, in the fourth and fifth row the color stretch is 10 times smaller for the spin-1 and spin-2 windows than in the first three rows.

becoming however increasingly more important. Nevertheless, the dominant B-mode contribution on subdegree scales is due to the gravitational lensing of the E-mode polarization and knowledge of the latter should again be sufficient. We elaborate more on these observations in the next Section.
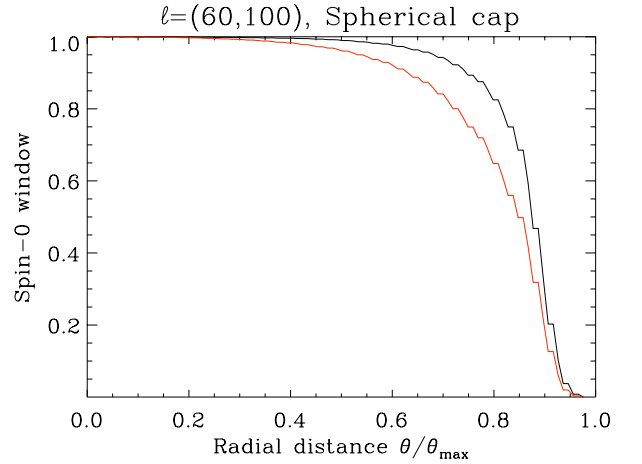

FIG. 6: Radial profile of the harmonic domain optimized spin-0 window function for the noise dominated case (red curve) and the signal-dominated case (black curve). For the signal-dominated case, the primodial B-mode has been set to $T/S = 0.5$ to emphasize the discrepancy between the two cases, all the other parameters are set as in Fig. 5.

Nonetheless as the prior assumptions have an impact on the final results they need to be chosen with care. In the case of the pixel-domain optimization we illustrate the dependence of the recovered windows on the assumed sky signal in Fig. 5 considering five different assumptions about the signal, while always keeping the noise level fixed at $\sigma = 5.75$ μK-arcmin. In the first three cases we assume the E mode power as in the best-fit WMAP 5-year model and change only the B signal. The three cases correspond to (i) no B-mode, (ii) only lensing-induced B-mode and (iii) both lensing-induced and primordial B-mode for $T/S = 0.05$. In the two remaining cases we fix the B-mode power as in the case (iii) and change that of the E-mode to be: (iv) a hundredth of the standard value and (v) zero. In all the cases, shown in Fig. 5 from top to bottom, the windows have been optimized for $\ell \in (20, 60)$ where a high amount of leakage has to be corrected for. For assumptions (i) and (ii), the B signal is lower than the noise for all angular scales whereas it exceeds the noise level in the last three cases. The shape of the optimized windows slowly varies with the B-mode level meaning that the E/B leakage from the dominant E-mode is the main source of extra variance and the optimization procedure focuses on resolving this particular source. In the last two cases, the E-mode is first set to be comparable to (case iv) and then lower (case v) than the B-mode power, what causes clearly discernible changes in the shape of the computed apodizations. In the next Section we discuss consequences of the observed changes in the window functions on the recovered power spectra variance.

In the case of the harmonic window the optimized weight functions depend on the noise and B-mode signal prior only, as emphasized by Eq. (29). The B-mode

power spectrum appears explicitly in $C^{\chi}_{\ell,\ell''}$ as well as in μ and λ. Unlike in the pixel-domain case, the E-mode signal is not involved here as by design the harmonic window perfectly fulfills the Dirichlet + Neumann boundary conditions, ensuring that the E-mode power leaked to B vanishes up to pixel effects. The dependence of such a window on the noise and B-mode prior is depicted on Fig. 6. The figure displays the radial profile of the spin-0 window in the noise-dominated (red curve) or the signal-dominated case (black curve). In the noise-dominated case, the power spectrum in $C^{\chi}_{\ell,\ell''}$ behaves as ℓ'^4 and our numerical result coincides with the analytic expression given in Eq. (33) of [16] applied to a spherical cap as the survey region, small enough for the flat sky approximation used in [16] to be valid. The apodization length of the window in the signal-dominated case is smaller due to the fact that in this case $C^{B}_{\ell} \propto \ell^{-2}$ and thus the power spectrum in $C^{\chi}_{\ell,\ell''}$ behaves roughly like ℓ'^2. This conclusion corroborates the analytic investigation of [16].

B. Analytic apodizations

Analytic window functions are in general not well suited for providing an optimal performance in the presence of the instrumental noise, given their limited ability to adapt to the specific properties of the data. However, they are quick to calculate, do not require any prior knowledge about the sky signals and can fulfill the boundary conditions with high precision. They may therefore at least be useful on exploratory, initial stages of any CMB data analyzes. Moreover, whenever prior information about the sky signals is available, the parameters of the analytic window functions can be tuned via Monte Carlo (MC) simulations providing a competitive performance to the optimized windows usually at much lower costs. The analytic windows are therefore of considerable practical interest and in this Section we will discuss two specific proposals of the analytic windows, elaborating and complementing the discussion presented in [15].

First of the considered choices is a straightforward generalization of the analytic window function proposed in [15] extended to be applicable to an arbitrary sky coverage. Calling δ_i the distance between the i-th observed pixel and the boundary, i.e., the smallest angular distance form the considered pixel to the contour of the mask, we can write the relevant expression as,

$$W_i = \begin{cases} \frac{1}{2} - \frac{1}{2}\cos\left(-\pi\frac{\delta_i}{\delta_c}\right), & \delta_i < \delta_c, \\ 1, & \delta_i > \delta_c, \end{cases} \quad (30)$$

and is zero outside the mask. Here, δ_c is a width of a boundary layer of the window smoothly interpolating between the window core and the zero on the outside. In the following, we refer to it as an apodization length and treat it as an adjustable parameter. Hereafter we
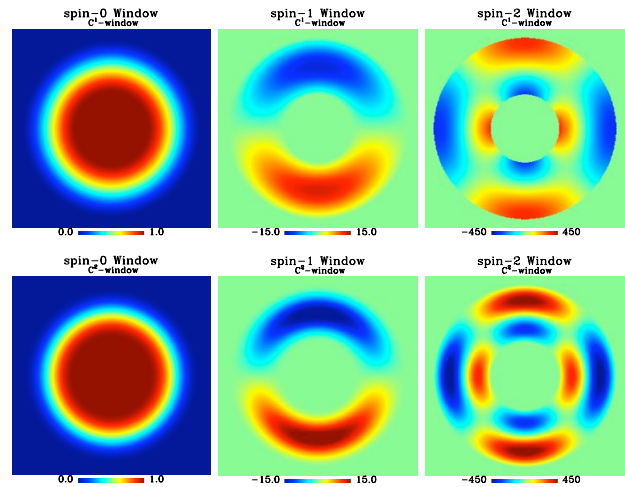

FIG. 7: Analytic window functions, spin-0, spin-1 and spin-2, left-to-right, respectively, calculated using the C^1 (upper panels) and C^2 (lower panels) functions. Only real parts of the spin-1 and 2 windows are shown.

will refer to the window defined in Eq. (30) as the C^1-window. One can easily check that if $\delta_i = 0$ (i.e., the pixel is on the contour of the mask), the window function is indeed vanishing, as well as that spin-1 window which can be derived from it. Therefore, the window defined above indeed fulfills the conditions as required by the pure pseudo power spectrum formalism. We note however that the related spin-2 window is not apodized and potentially therefore could be a source of numerical problems in particular, whenever SHT are involved.

We therefore propose a second analytic formula for the spin-0 window from which apodized spin-1 and spin-2 windows can be derived, reducing in principle the ringing due to a sharp spin-2 window. With the same notation as above and for every pixel in the mask the window is defined as,

$$W_i = \begin{cases} -\frac{1}{2\pi}\sin\left(2\pi\frac{\delta_i}{\delta_c}\right) - \frac{\delta_i}{\delta_c}, & \delta_i < \delta_c, \\ 1, & \delta_i > \delta_c, \end{cases} \quad (31)$$

and vanishes outside the mask. We will refer to this choice as the C^2-window.

The three spin windows, for each of the two cases, are shown in Fig. 7 in a case of a spherical cap. The apodization length is set to 7 degrees. The spin-0 window indeed falls more rapidly to zero in the second proposal, thus leading to an apodized spin-2 window. Nevertheless the two analytic windows are very similar with only small differences concentrated around the edge of the observed patch. We will discuss an impact of such differences in the window shapes on the E/B leakage in the next Section.

Two remarks are in order here. First of all, the analytic windows can account for the presence of holes in the mask for example due to masked out point-sources or unobserved pixels. The distance δ_i just has to be computed
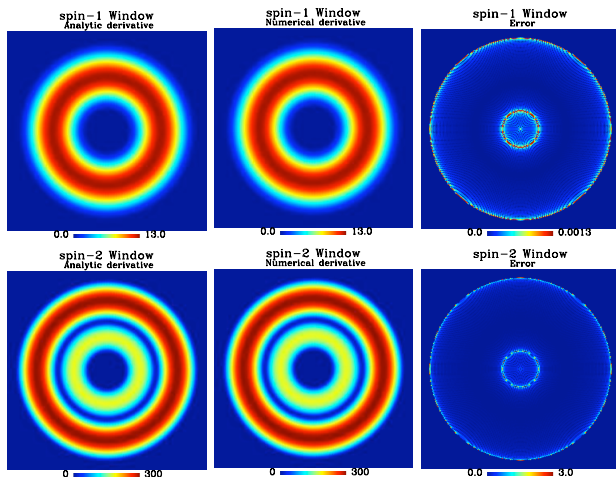
FIG. 8: Absolute value of the analytic spin-1 windows (upper panels) and spin-2 windows (lower panels) derived via the analytic (left line) and numerical computation (middle line). The plots in the right panels show the respective differences. Note that in the right panels the adopted color scale spans a range of values 10^{-4} (up) and 10^{-2} (bottom) times smaller than in the remaining panels.

between the i-th pixel and the boundary closest to it (external or internal). Second, for an arbitrary sky coverage (including holes), the spin windows need to be clearly derived numerically, however, in some simple cases such as that of a spherical cap or square patches the spin-1 and spin-2 windows can be computed analytically from the spin-0 window. We can, and will, use that latter observation to quantify some numerical effects, involved in the discussed formalism. For instance we will investigate the effects of the numerical derivatives on the shape of the spin-weighted windows, as shown for a spherical cap case in Fig. 8.

IV.   TESTS AND APPLICATIONS

Performance of the technique discussed here depends on many factors, ranging from the pixelization effects to assumed knowledge of the sky signals, and can be looked at from different perspectives. In this Section we discuss some of the most relevant effects. As the presented estimator first aims at solving the E/B leakage, we start this Section with a discussion of a quality of the E/B leakage control and follow on with a study of the estimator efficiency expressed in terms of the B-mode power spectrum variance levels, which can be achieved using this approach in different circumstances. Unless it is explicitly mentioned otherwise, we will work with the HEALPix pixelization scheme adopting the resolution $N_{side} = 512$. The input E-mode signal is that of the cosmological model with parameters as constrained by the WMAP 5-year data [36]. The B-mode signal includes lensing and primordial B-mode with $T/S = 0.05$.

For the case of homogeneous noise the noise level is set to 5.75 μK-arcmin, which corresponds to a typical level expected for the small-scale B-mode experiments.

To study the dependence on the patch geometry, we adopt five different patch shapes: a spherical cap, a square and three rectangular patches with a different elongation. Their geometrical properties are summarized in Table I. All the patches cover the same sky area (roughly 1%) but the length of their contour increases progressively from the spherical cap (C) to the most elongated rectangle (R3).

		θ [deg.]	φ [deg.]	perimeter [rad.]
SQUARE	(S)	20	20	1.38
RECTANGLE 1	(R1)	16	24.9	1.42
RECTANGLE 2	(R2)	12	33.2	1.57
RECTANGLE 3	(R3)	8	49.8	2.01
CAP	(C)	radius=11.3 deg.		1.23

TABLE I: Geometrical properties of the five patches used in the examples considered here: θ is the zenithal aperture, φ the azimuthal aperture and the perimeter is the total geometrical length of the contour of the patch. The five patches cover the same sky area roughly equal to 1% of the entire celestial sphere.

A.   Mode mixing kernel

The mode mixing kernel provides the most direct measure of the E/B leakage. The root-mean-square (rms) of the fraction of the power contained in one of the two polarized modes which is leaked to the other mode is characterized by the magnitude of the off-diagonal elements of the mixing kernel. For the pure estimators in principle those should be vanishing if the proper apodization window is used. However, in practice, the apodization is only one of the factors on which the level of the leakage depends and numerical effects due to approximations and simplifications involved in the calculations of the estimated spectra give often rise to the power leakages. In this Section we study dependence of the off-diagonal blocks of the mixing kernels on the specific choices of the apodization windows, pixelization scheme, pixel size, sky location and shape of observed sky areas, as well as errors due to the approximations involved in the numerical derivative calculations.

We show an example of the mixing kernels for both standard and pure pseudo-spectra in Fig. 9. In the standard case we note a strongly band diagonal character of both, diagonal and off-diagonal, blocks of the mixing kernel with a magnitude of the diagonal elements of the off-diagonal block comprising a significant fraction of the corresponding elements in the diagonal block. This is a clear indication of the leakage. The off-diagonal block elements are manifestly suppressed in the case of the pure
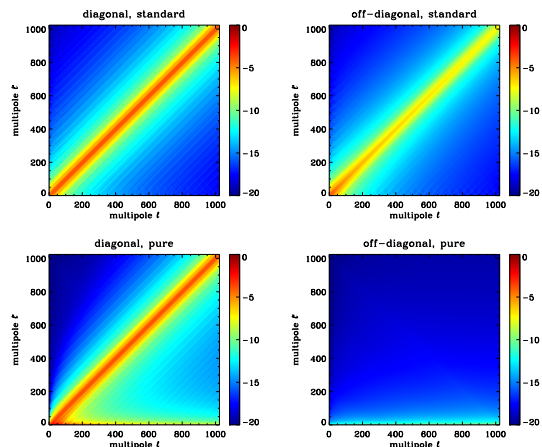
FIG. 9: *Upper row:* Diagonal, BB, (*left*) and off-diagonal, EB, (*right*) blocks of the mixing kernels for the standard estimations. *Lower row:* Same as above but for the pure estimator computed using the analytic C^2-window. The color scale is logarithmic and the same in all the panels. The vertical axes correspond to the power spectrum multipoles, while the horizontal ones to the pseudo-spectrum, Eq. (12).

estimator, as expected given that in the displayed case we have used a properly apodized, analytic C^2-window. Also as expected the diagonal block remains strongly diagonal dominated. Nevertheless there is a substantial level of mixing between the low-ℓ multipoles of the power spectra with essentially all multipoles of the corresponding pure pseudo-spectra seen for both the blocks. This extra mixing arises due to the presence of the counterterms in the definition of the pure modes, Eq. (24). As a consequence a special treatment on the map level aiming at reducing the power contained in the modes comparable and larger than the observed sky patch may be required prior to the power spectrum estimation.

1. Pixel size

As we already hinted at before (Sect. II D 2), pixelization of the sphere is a source of residual E/B leakage, which will in general depend on the pixel sizes and shapes. The dependence of such a leakage on the HEALPix pixel size for two types of window functions is displayed in Fig. 10. Each of the three lower curves shows a single column of the off-diagonal block of the mixing kernel, $M_{\ell\ell'}^{off}$, computed using the C^2-window and corresponding to, from top to bottom, $N_{side} = 256$, 512 & 1024. The residual, leaked E-mode decreases for higher resolution confirming the pixel origin of this leakage. The three upper curves in Fig. 10 show the case of the optimized window computed in the pixel-domain. These windows have been optimized for the range of ℓ from 60 to 100 and we kept the noise per pixel area constant. In all three cases we find a substantial level of the residual leakage, which is independent on the adopted pixel size.

This leakage appears therefore mostly due to the cut-sky effects, which are, by design, not fully removed in the case of the optimized windows which do not fulfill the proper boundary requirements, and are only suppressed to the levels below that of the other uncertainties.
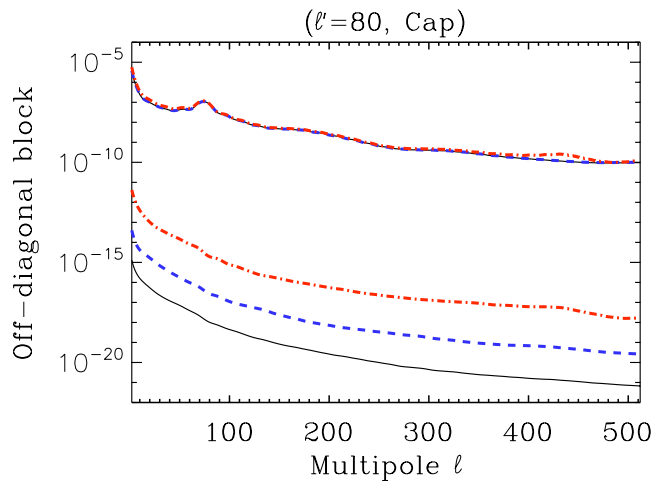

FIG. 10: Off-diagonal block of the mixing kernel $M_{\ell\ell'}^{off}$ for two different classes of windows and for different HEALPix pixel sizes: $N_{side} = 1024$ (black-solid curves), $N_{side} = 512$ (blue-dashed curves) and $N_{side} = 256$ (red-dashed-dotted curves). The three lower curves are for the C^2-window and the three, overlapping upper curves – for the optimized window computed in the pixel-domain. The apodization length has been set equal to 9 degrees for the C^2-window and the optimized windows has been optimized for $\ell \in (60, 100)$. For the analytic windows, the residual leaked E-mode is induced by pixel effects and therefore decreases for the higher resolution. For the optimized window, the residual leakage is mainly due to a sky cut and as such does not depend strongly on the pixel size.

2. Window choice

A comparison of the residual levels of the E-to-B leakage for the different sky apodizations as considered here is presented in Fig. 11. We plot there a single column of the off-diagonal block of the mixing kernel for $\ell' = 80$ and computed for the four types of windows defined in Sect. III and the spherical cap area as defined in Table I. As expected the optimized window computed in the pixel-domain leads to the highest amount of leakage. The harmonic space window fares better in part thanks to the boundary conditions which are explicitly imposed on it and in part due to the extra-apodization. For the two types of optimized windows, the level of leakage varies slightly depending on a prior assumed for the B-mode power.

Both the analytic windows satisfy the Dirichlet and Neumann conditions and therefore lead only to a residual amount of E-to-B leakage. Nevertheless, the C^2-window
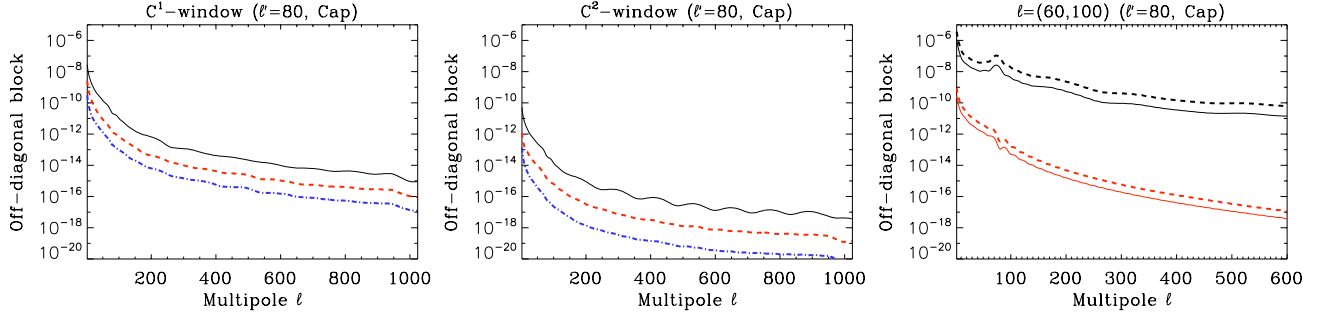
FIG. 11: A $\ell' = 80$ column of an off-diagonal block of the mixing kernel $M_{\ell\ell'}^{off}$ for the four classes of apodized windows. *Left panel :* C^1-analytic window for three different values of the apodization length : $\delta_c = 3$ deg. (black-solid curve), $\delta_c = 5$ deg. (red-dashed curve) and $\delta_c = 8$ deg. (blue-dashed-dotted curve). *Middle panel :* Same as left panel but for the C^2-window. *Right panel :* optimized windows computed in the pixel-domain (black curves) and in the harmonic domain (red curves). Solid curves are for optimized windows derived without B-mode and dashed curves for optimized windows with lensing-induced and primordial B-mode for $T/S = 0.05$.

performance is superior to that of the C^1-window, producing the level of leaked E-mode 10^{-4} to 10^{-2} smaller. This is because the level of the leakage is reduced by increasing the apodization length. For the analytic windows, more apodization is achieved either by increasing the apodization length or by additionally forcing the spin-2 window to vanish at the edges of the survey. For the optimized windows, as discussed before the amplitude of the adopted B-mode prior leads to an increase of the effective apodization length, thus affecting the leakage level. Because of pixelization, the windows are discretized and the Dirichlet and Neumann boundary conditions are inevitably broken. By increasing the apodization, we reduce the size of the discrete jump in the window values between two neighbouring pixels at the patch boundary, thus mitigating the extent to which that affects the boundary conditions.

3. Derivatives

For arbitrary patch geometries and in the case of the analytic apodizations the spin-1 and spin-2 windows may have to be computed from its spin-0 counterpart numerically and therefore being often only approximate. In this paper, as described before, we perform the numerical derivation in the harmonic domain. The effect of using the approximate rather than exact nonzero spin windows is depicted in Fig. 12, where again a single column of the mixing kernel is shown for the spherical cap computed using either the analytic formula (in black) or numerical computations (in red) for spin-1 and spin-2 component. This shows that calculating the derivatives numerically does not affect the diagonal block elements and leads only to a slight increase of the E-to-B leakage, as expressed by a small change in the off-diagonal block element magnitudes. This latter effect is indeed very small and only appears at the smallest scales, confirming thus that our prescription for the derivative computation permits us to
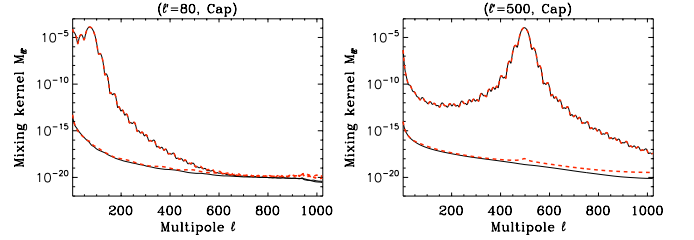

FIG. 12: A column of a diagonal (upper curves) and off-diagonal (lower curves) block of the mixing kernel using analytically (black-solid curves) and numerically (red-dashed curves) computed spin-1 and spin-2 windows. The dashed and solid curves overlap over most of the angular scale. Left panel shows a column calculated for $\ell' = 80$ and right – for $\ell' = 500$. The numerical computation of the spin-1 and spin-2 windows leads to a slight increase of the leakage in the high ℓ part seen as a difference in the off-diagonal block curves.

retain very good control of the leakage, especially in the low-ℓ part, where it is of primary importance.

4. Patch geometry

The remaining leakage as a function of the geometry of the observed sky area is depicted in Fig. 13 for the two analytic windows and for two values of the apodization length. For an apodization length of 3 degree the level of leaked E-mode is the same in all the considered cases, demonstrating that we can achieve a good control of the leakage for any elongation. For $\delta_c = 5$ degree the leakage for the most elongated patch, i.e., the R3 geometry, is greatly enhanced. This is because for such a patch, apodization lengths greater than 4 degree exceed the half of its smallest size causing the spin-0 window, as defined in Eqs. (30) or (31), to be strictly speaking nondifferentiable. Nevertheless the spin-1 and spin-2 windows can still be formally introduced, using the fact that one-sided

derivatives exist. That however unavoidably results in the spin-1 and spin-2 windows, which are not always continuous, leading the leakage excess. This emphasizes the importance of ensuring that all three spin-s windows are continuous throughout the entire observed region – a condition, which on occasion may require either trimming some irregular edges of realistic sky patches as observed by actual experiments or applying an extra smoothing to the spin-0 window. Nevertheless, whenever this condition is fulfilled the leakage is largely independent on the observed patch shape, and with the help of the presented formalism can be, at least in principle, removed down to the level as defined by the pixelization effects.
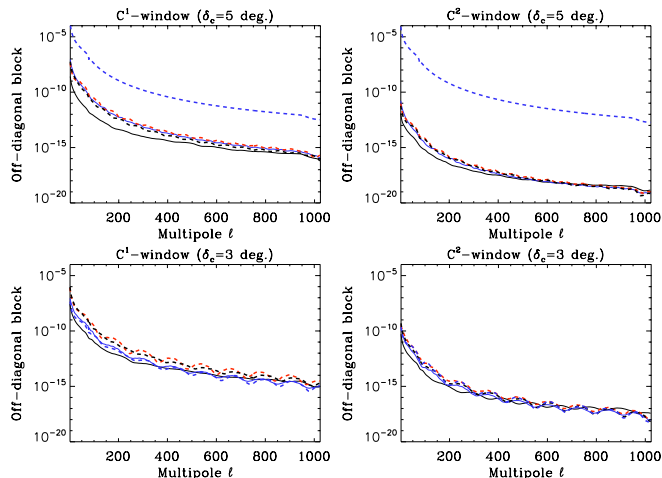
FIG. 13: A column of the off-diagonal block of the mixing kernel $M^{off}_{\ell\ell'}$ for $\ell' = 80$ and for five different patch geometry: spherical cap (black-solid curve), square (blue-solid curve), rectangle R1 (red-dashed curve), rectangle R2 (black-dashed-dotted curve) and rectangle R3 (blue-dashed curve). The left panels show the results for the C^1-window and the right panel for the C^2-window. Upper panels are for $\delta_c = 5$ deg. and lower panels for $\delta_c = 3$ deg. For the (R3) geometry, an apodization length of 5 deg. exceeds the half of the smallest size of the patch. As a consequence, the spin-1 and spin-2 windows are not continuous right in the middle of the patch, leading to a significant increase of the pixel-induced leaked E-mode (see blue-dashed curves of the upper panels). For all the other cases, the curves overlap proving the good control of the E/B leakge irrespectively of the geometry.

The case of optimized windows computed in the pixel-domain is displayed on Fig. 14 where the level of leakage is shown for the windows optimized in the $\ell \in (60, 100)$ bin. The noise and signal prior is the same for all the patches. As for the analytical sky apodizations, the level of leakage only marginally depends on the patch geometry. This means that the optimization procedure manage to lower the level of leaked E mode below the noise irrespectively of the shape of the boundaries (we remind that the five surveys cover the same sky area, which roughly leads to the same amount of noise in ℓ-space). We note that as expected the level of leakage allowed by the optimized windows is a few orders of magnitude higher than
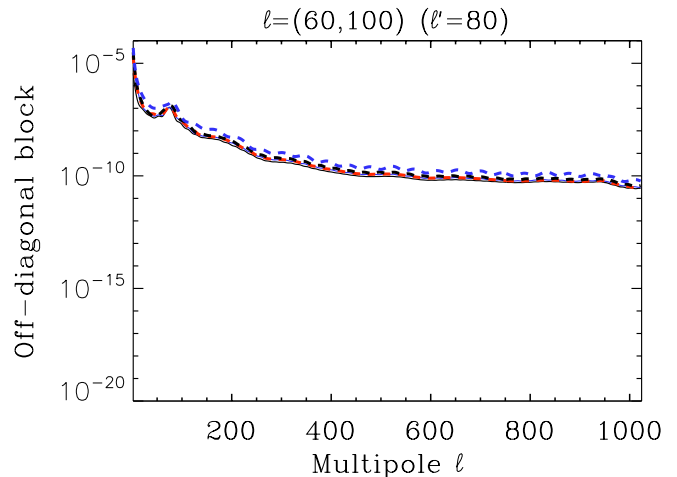

FIG. 14: A column of the off-diagonal block of the mixing kernel $M^{off}_{\ell\ell'}$ for $\ell' = 80$ and calculated for windows optimized in a pixel-domain for a range $\ell \in (60, 100)$ and the five different patch geometries (**the five curves overlap**): spherical cap (black-solid curve), square (blue-solid curve), rectangle R1 (red-dashed curve), rectangle R2 (black-dashed-dotted curve) and rectangle R3 (blue-dashed curve). The axis ranges coincide with those in Fig. 13.

for the analytic windows.

5. Dependence on a sky position/pixelization

Numerical computations of derivatives of spin-0 windows will depend in general on the sky location and on the adopted pixelization scheme, even if the windows are defined analytically. For instance, in the case of a spherical cap centered at the north pole, the precision of a calculation of the azimuthal part of the SHT quickly deteriorates in the HEALPix scheme because of the decreasing number of constant declination pixels present in the polar areas. As a consequence, the spin-1 and spin-2 windows computed in the harmonic space can, sometimes drastically, depart from their true behavior often exaggerating the level of the leakage. Such a problem can be cured either by rotating the cap close to the equatorial plane, as often plausible in cases with a small sky coverage, or by using a different pixelization/gridding scheme such as GLCP (i.e., Gauss-Legendre Celestial Pixelization – ECP-like but with the constant declination rows distributed as the zeros of the Legendre polynomial of some order as supported by the S^{2}HAT software [30]), for which the number of azimuthal pixels (grid-points) remains the same for each constant declination row, and which provide a nearly exact quadrature over the polar angle. We provide examples of the potential effects in Fig. 15. We note that the choice of the pixelization used for the derivative computations can be different from that used for the analyzed maps, though care needs to be taken while performing interpolations required in such
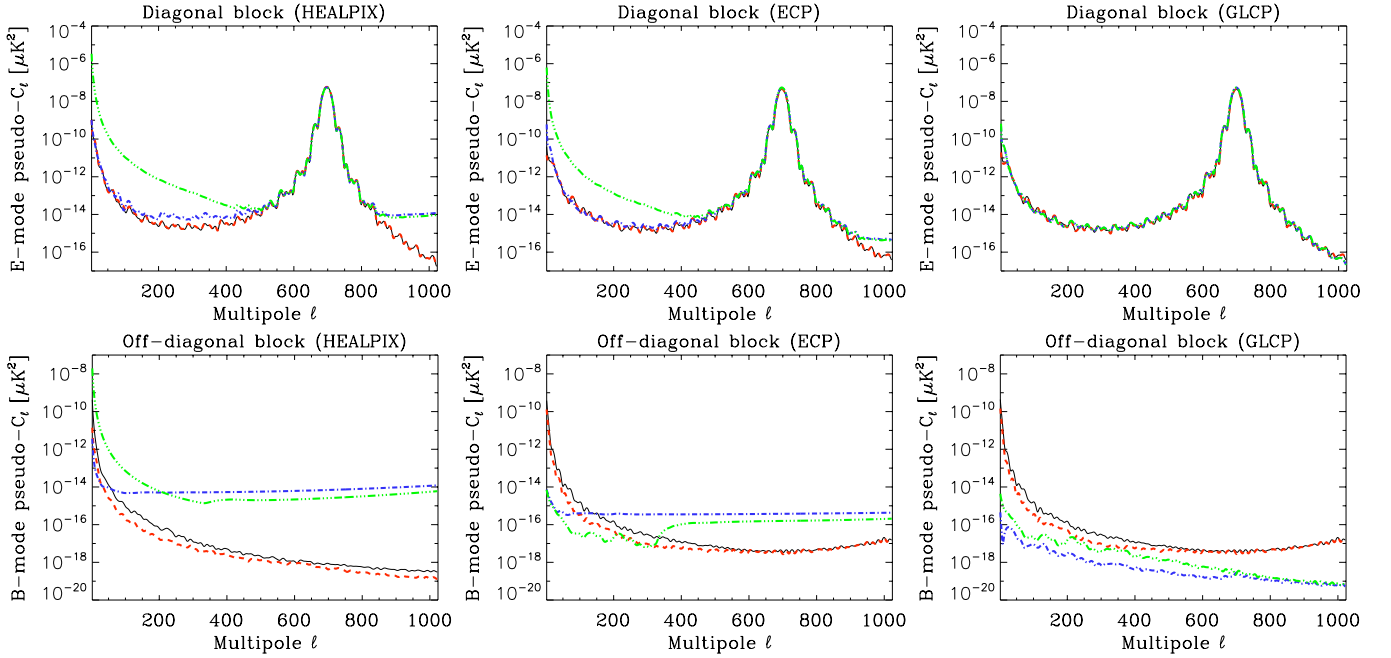
FIG. 15: An impact of the sky location and the pixelization scheme on the diagonal block (upper panels) and off-diagonal block (lower panels) of the mixing kernel $M_{\ell\ell'}$. The different curves partially overlap for the diagonal block. *Left panels :* Computation in the HEALPix scheme for a cap centered at the equatorial plane using analytic spin-1 and spin-2 windows (black-solid curve) or numerical spin-1 and spin-2 windows (red-dashed curve) and for a cap centered at the north pole using analytic spin-1 and spin-2 windows (blue-dashed-dotted curve) or numerical spin-1 and spin-2 windows (green-dashed-dotted curve). *Middle panels :* Same as the left panels for ECP pixelization scheme. *Right panels :* Same as the left panels for GLCP pixelization scheme. The window function used here is the C^1-window with the apodization length set $\delta_c = 8$ degrees. In all three cases the pixelization parameters are chosen in such a way that there is the identical number of the constant declination rings within the the sky cap for each of them. For the HEALPix scheme, the bad sampling of the sphere close to the north pole (particularly along the azimuthal direction) leads to strong pixelization effects becoming dominant even for the diagonal block. A similar problem is also seen in the ECP case, but is resolved in the case of the GLCP scheme, which properly samples the sphere all the way to the poles in both polar and azimuthal directions. The effect can be also alleviated by shifting the cap closer to the equatorial plane.

a case. In addition, similar effects may also play a role whenever any SHT is to be done, as for example, in the computation of the pure multipoles, and therefore choice of the pixelization and/or its orientation are of primary importance for techniques as the one discussed here.

B. Power spectra uncertainty

The errors of the estimated B-mode power spectra include contributions due to both the signal and the noise. For any ℓ-bin the signal sample variance is due to the variance of the B-mode signal modes from that bin but also from other multipoles aliased to it. It is also due to power of the leaked E-mode into B, i.e., the E/B leakage. The window functions discussed before lead to different trade-offs between these different sources of the uncertainties. In this Section, we will assess the power spectra uncertainty using MC simulations comparing the performances of the pure pseudo cross power spectrum approach for different choices of the window functions. We will also consider the effects due to patch geometry

and signal priors, complementing the discussion in the previous Section.

Hereafter, we estimate the variance as the standard deviation of 1,000 MC simulations. Unless specified explicitly otherwise we will employ the same set of parameters as before, i.e., we will work with the HEALPix scheme with the resolution $N_{side} = 512$ and assuming the WMAP 5-year data cosmological model [36] for the input E-mode signal. The B-mode signal includes lensing and primordial B-mode with $T/S = 0.05$. We first focus on the homogeneous noise case with a level of 5.75 μK-arcmin. The estimated power spectrum are binned with the bins of a width of $\Delta\ell = 40$ and with the lowest-ℓ bin starting at $\ell = 20$. Except for the cases investigating effects of the patch geometry, the simulated maps cover a spherical cap (C) area centered at the equator as described in Table I.

As a benchmark for the derived numerical results we will use the simplified, theoretical variance derived for a
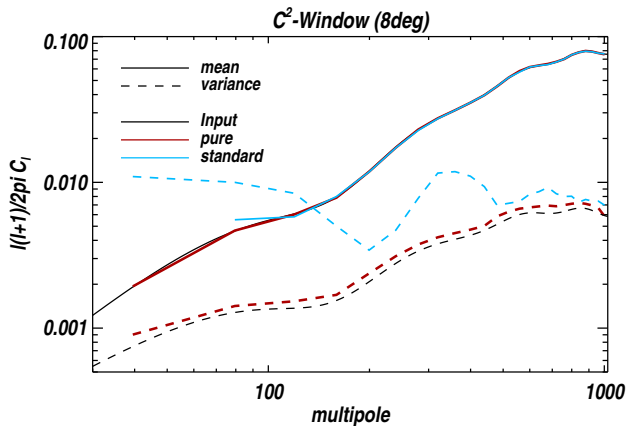
FIG. 16: Average B-mode spectra (solid curves) and their sample variances (dashed curves) obtained using standard, in blue, or pure estimators, in red, and computed from 1000 signal-only MC simulations done on a spherical cap covering 1% of the sky. Also shown are the input power spectrum (black-solid) and the Fisher estimate of the variance (black-dashed). The extra-variance for the standard estimators comes from the E-mode power leaked into B, while the sample variance for the pure estimators is close to optimal.

cross-spectra from the Fisher consideration [11, 29],

$$\Delta \mathcal{C}_\ell^B = \sqrt{\frac{2 f_{sky}^{-1}}{(2\ell+1)} \left(C_\ell^{B\,2} + C_\ell^B \frac{\sigma^2}{B_\ell^2} + \frac{\sigma^4}{B_\ell^4} \right)}, \qquad (32)$$

where σ^2 is the noise power spectrum and f_{sky} will be hereafter taken to be an actual covered sky area. As such this uncertainty provides a lower limit on any power spectrum estimation error, including the one derived using the pure estimator discussed here, and therefore a useful reference to gauge the performance of the pure method against. We emphasize that we do not expect that our estimator will be ever able to reach this limit even in the cases when no E/B leakage is allowed for, i.e., when C_ℓ^E is set to 0. In fact a tighter lower limit could be readily obtained, if an effective, and thus usually smaller, rather than actual sky area is used as f_{sky} in Eq. (32) accounting for the presence of the apodizations in the pure formalism. However, even then the excess variance is usually seen owing to the presence of the counter terms in Eq. (24).

We illustrate some of these effects in Fig. 16, where we present the results of the signal-only MC simulations for both standard and pure pseudo-spectra approaches and contrast them with the semianalytic predictions as calculated using the Fisher approach, Eq. (32).

1. Effect of the window function

The variance of the estimated B-mode power spectra for a number of the apodization lengths and the two analytic proposals is displayed in Fig. 17 and contrasted against the variance derived using the optimized windows computed in the pixel-domain (red curve) and its Fisher estimate (black-dashed curve). As a reference, the input theoretical B-mode used in the simulations is plotted as black-solid lines. Given the adopted here sky coverage (and hence bin-width) as well as the noise level, both in a ball-park of the anticipated experiments, we find that in the range of ℓ from 20 to 1000 the variance is always few times smaller than the signal itself thus in principle allowing for a good reconstruction of the B-mode power spectrum.

The apodization length has clearly an effect on the performance of the method. For the C^1-windows, the lowest variance is reached using a 5 degree apodization in the first bin and 2 degrees in all the others. For the C^2-windows, 4 and 3 degrees should be used in the first and second bins respectively, whereas the 2 degree window gives the lowest variance for all the other bins. This is due to the fact that noise in the subsequent bins is bigger for the higher ℓ values and thus though the leakage level somewhat increases for the smaller apodizations, boosting up the variance. This is compensated by the simultaneous decrease of the noise variance. At least for the homogeneous noise the best optimization length can be clearly tuned for every analytic window and each bin if some prior information about the sky power is available. This can be done determining via MC simulations the total estimator variance for a range of considered apodizations and selecting the one leading to the minimal scatter.

The variance derived using the optimized windows computed either in the pixel-domain (red curve) or in harmonic space (blue curve) is displayed in Fig. 18. The two windows lead to a similar level of variance, which is also on par with the Fisher forecast at least for $\ell > 200$. Clearly, in the cases studied here the use of pure pseudo-spectrum technique yields results close to the best achievable for angular scales smaller than a degree. Moreover, the derived variance is indeed at least no larger than the one obtained with the analytical windows as can be seen in Fig. 17. The very good agreement between the harmonic space computation and the pixel-domain computation of the optimized windows shows that the extra apodization used during the harmonic computation does not cause any significant departure from optimality. This justifies the use of the harmonic window in the homogeneous, or nearly so, noise cases as it is much faster, and better controls the E/B leakage across the entire range of ℓ and does it without any apparent loss of performance for the final precision of the estimated power spectrum.

2. Effect of the signal prior

As we have emphasized before the optimization procedure requires knowledge of both the noise and signal properties, including the B-mode. In most of the cases of the forthcoming CMB polarization experiments noise
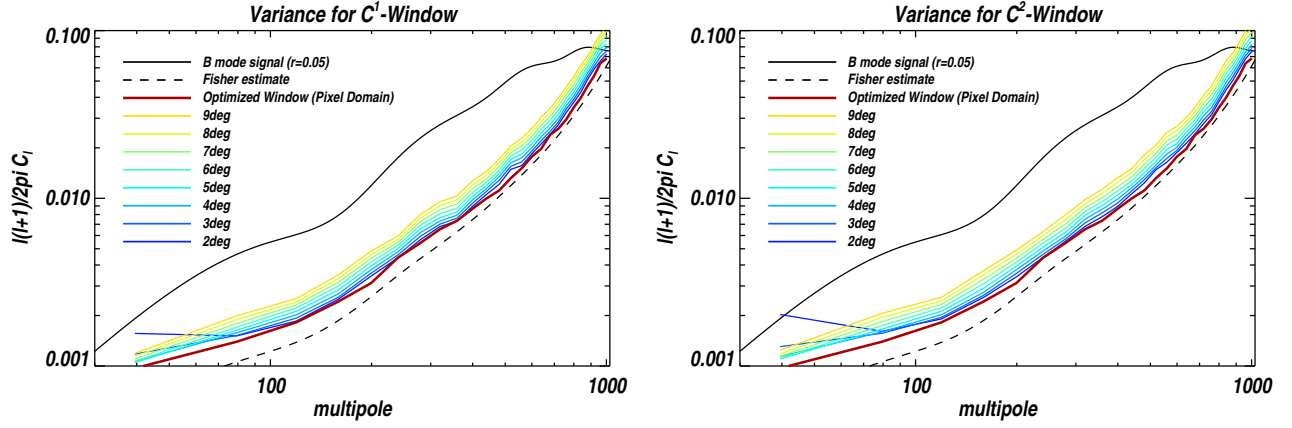
FIG. 17: Variance of the estimated B-mode power spectrum using the C^1-window, *left panel*, and the C^2-window, *right panel*, for different apodization lengths: from 2 (blue) to 9 (yellow). The red curve shows the variance obtained using the optimized windows computed in the pixel-domain and optimized bin-by-bin. The observed survey is the spherical cap (C) and the cosmological model and the level of homogeneous noise are described in the text. The input theoretical B-mode is shown with a black-solid line and a Fisher estimate of the variance with a black-dashed curve.
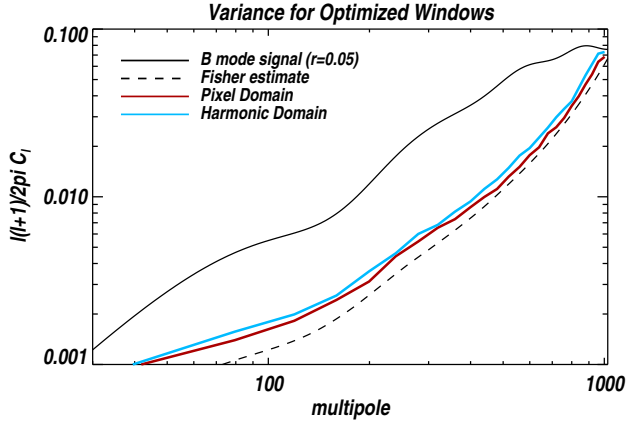

FIG. 18: Variance of the estimated B-mode power spectrum using the optimized windows computed in the pixel (blue) and harmonic space (red curve). The black-dashed curve shows the Fisher estimate of the variance and the black-solid line – the input theoretical B-mode. The same performance is achieved using the two implementations of the optimized window and is close to optimal for $\ell > 200$.

and E-mode properties will be probably sufficiently constrained for high and intermediate ℓ modes, so the major source of uncertainties will result from our ignorance of the primordial B-mode and low-ℓ E-mode power.

To assess the influence of the B-mode prior on the performance of the optimized windows, we have performed simulations using lensing-induced and primordial B-mode ($T/S = 0.05$) as input signal but estimating the power spectrum using three different sets of windows optimized for different B-mode signal : (1) no B-mode, (2) lensing-induced B-mode and (3) lensing-induced plus primordial B-mode for $T/S = 0.05$. As the two first assumptions differ from the B-mode used to synthesize the map in the simulations, we expect that the optimized

windows derived under those hypotheses could occur to be suboptimal. The results of such simulations are shown in Fig. 19, left panel, where the signal variance (solid curves) and the noise variance (dashed curves) for the three sets of optimized windows are displayed. The black curves correspond to their Fisher estimates. It clearly shows that the prior on the B-mode used in the optimization procedure does not significantly affect the performance of the pure pseudo-spectrum estimation, as long as the postulated power of the B-mode remains much smaller than that of assumed in the E-mode prior. This last condition will be always fulfilled taking into account the current upper limits on the B-mode at least as long as CMB is the only sky signal considered. As a consequence, a conservative, but sufficient, i.e., nearly optimal approach would consist in assuming lensing-induced B-mode for optimization, as its level and shape will be largely known from the modeling of the E-mode spectrum.

Though in general we expect to have a sufficiently well constrained E-mode power spectrum for the future polarization-sensitive experiments, that may not be always so at least as far as the very lowest multipoles are considered. To test the impact of those on the variance of the recovered power spectra we have computed the optimized windows with a fixed B-mode power spectrum, including lensing-induced and primordial B-mode for $T/S = 0.05$, but using three different E-mode priors. The three E-mode power spectra used in this test were calculated for three different values of the reionization optical depth τ: (1) $\tau = 0$, (2) $\tau = 0.2$ and (3) $\tau = \tau_{WMAP} = 0.087$ as constrained by the WMAP 5-year results. These three power spectra have been used to first compute the corresponding optimized windows and later to perform MC simulations in order to obtain the variance estimates. The 'true' sky used to produce the maps to-be-analyzed here assumed $\tau = \tau_{WMAP}$ and
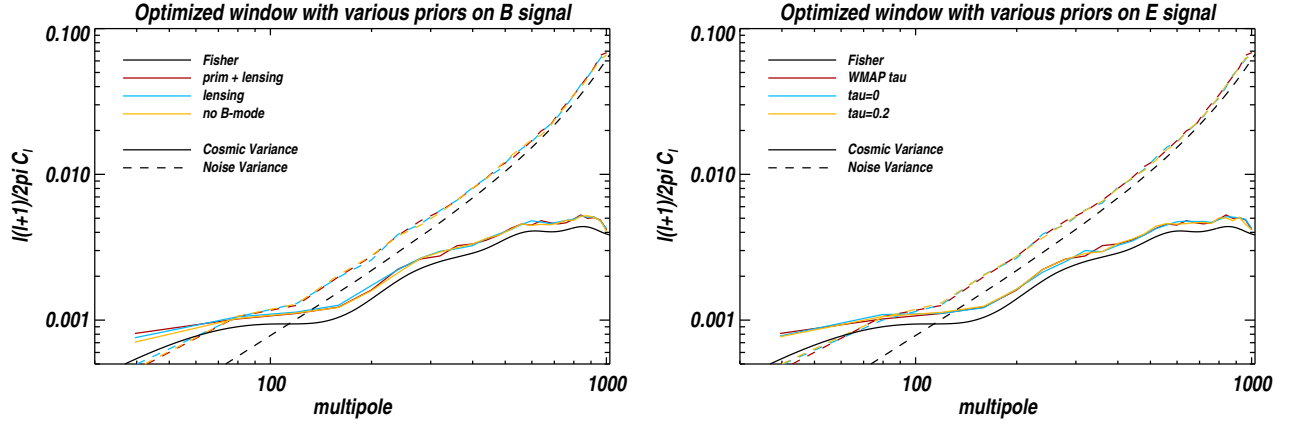
FIG. 19: *Left panel:* Noise (dashed curves) and sampling (solid) variance of the estimated B-mode power spectrum using pixel-domain optimized windows calculated for three different assumptions about the B-mode signal: no B-mode (blue curves), only lensing-induced B-mode (red curves) and lensing-induced plus primordial B-mode for $T/S = 0.05$ (yellow curves). The Fisher estimates of the noise and signal variances are shown in black. This is the 3rd case model, which has been used to generate the signal maps analyzed in all three cases. Though the assumed sky priors for B-modes in two first cases are wrong, that does not seem to have a strong impact on either the signal or the noise variance. *Right panel:* Same as left panel but using windows optimized for three different assumption of the E-mode signal via a varying reionization optical depth: $\tau = 0$ (red curves), $\tau = 0.2$ (yellow curves) and $\tau = \tau_{WMAP} = 0.087$ (blue curves). This last model was also used to generate the signal maps analyzed in all three cases. As before no strong impact of the priors on either the recovered signal or the noise variance is seen in any of the considered cases.

thus coincided with the third prior as considered above. The resulting noise (dashed-colored curves) and sample (solid-colored curves) for the three prior are displayed on Fig. 19, right panel, alongside the Fisher estimate (black-dashed curve) and the input C_ℓ^B (black-solid curve). We find that the variance is only marginally affected by the E-prior, even at large angular scales where a varying optical depth significantly changes the shape and amplitude of the E-mode. We note that our knowledge of the E-mode spectrum will be soon much more precise than the uncertainty permitted in the test described here, providing therefore an external prior certainly sufficiently accurate for the purpose of the method discussed in this paper.

3. Effect of the patch geometry

From the previous discussion, we have already learned that the shape of the observed sky area does not affect our ability to control the E/B leakage with the help of a proper apodization and pure formalism. We can therefore anticipate that the dependence of the variance of the estimated B-mode power spectrum on the patch geometry will follow the same pattern as for the standard pseudo-spectrum estimator.

We show the numerical results in Fig. 20, alongside the Fisher estimate, which is the same for all considered sky patches as they all have the same sky area (Table I). The power spectra have been estimated using the C^2-window function with an apodization length optimized for each bin and every patch separately. Except for the REC3
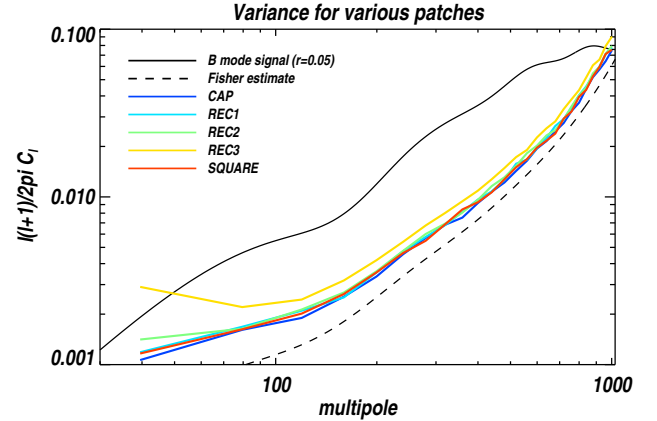

FIG. 20: Variance of the estimated B-mode power spectrum with respect to the patch geometry using the C^2-window with an apodization length optimized via MC simulations for each bin and each considered patch. The black-dashed curve shows the Fisher estimate of the variance and the black-solid line – the theoretical input B-mode spectrum. For $\ell > 200$, the estimation is close to optimal and a good reconstruction of the B-mode can be obtained. For smaller multipoles, the performance deteriorates for more elongated patches though a variance smaller than the expected power level is recovered in most of the studied cases.

patch, for which the variance of the estimated power spectrum exceeds the power level in the first bin, the variance is small enough to ensure a good reconstruction of the B-mode power spectrum. Moreover, for all the patches, it is roughly the same and on the level com-
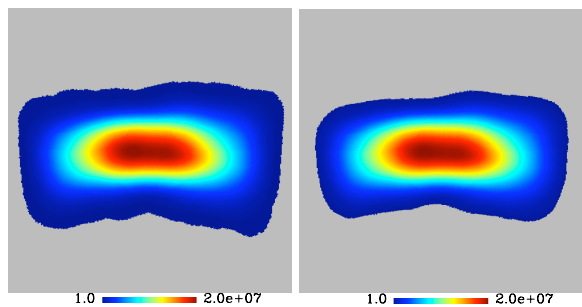
FIG. 21: A distribution of sky observation for an observational strategy mimicking that of the EBEX balloon-borne experiment. The left panel shows the full observed sky patch, while the right one – its well-observed part. The latter contains pixels for which a number of the observations is not smaller than 5×10^{-3} of the best observed pixel. The density of observations per pixel ranges from $\sim 2 \times 10^7$ in the center down to 1 or 10^5 at the edge in the full and reduced sky cases respectively. The pixel size is roughly 7 arcmin, corresponding to $N_{side} = 512$. The size of the shown panels is roughly $20^\circ \times 20^\circ$ each. The color stretch adopted here is linear.

parable to that of the Fisher variance at least down to $\ell \sim 200$. For larger angular scales however, the patch geometry starts affecting the performance of the estimation of the power spectrum as the variance increases for more elongated patches. A similar behavior can be seen also in the case of the optimized windows, as expected given that the patch geometry does not affect the level to which the E/B leakage is controlled.

C. More realistic example

The sky coverage of the realistic experiments is clearly more complex than any of the cases studied so far. The observed patches are usually irregular and a density of observations per sky area is pixel-dependent giving rise to significant noise inhomogeneity of sky maps recovered from data of such experiments. In this Section we therefore assess the performance of the pure pseudo-power-spectrum approach in the case of an observation mimicking a long-duration CMB balloon-borne experiment, such as for instance EBEX [28]. The adopted here sky coverage and its sampling, i.e., a number of observations per pixel, corresponds to one of the possible scanning strategies for that experiment and is shown in Fig. 21. As it is typical of the small-scale experiments the realistic distribution of the observations is very inhomogeneous. The number of samples per pixel ranges from 1 at the patch edge to roughly 2×10^7 in the center. Though in principle including all the pixels could be advantageous, if they are properly apodized, in practice it turns out not to be always the case. Moreover, we find that due the dynamical range for the noise levels the PCG-based, iterative calculation of the optimized windows fails to converge within a reasonable number of iterations, while a direct

inversion of the involved matrix is clearly too costly to be performed. For this reason in the following examples together with the full observed sky patch we consider also its well-observed part as shown in the right panel of Fig. 21. It contains all the pixels for which a number of observations is not smaller than 10^5, reducing the dynamical range of the noise levels from roughly 3000 to 15. This has an unavoidable and unfortunate impact on a observed sky area, which in the studied case decreases from 1.1% down to 0.89% of the total sky. We set the noise level per sample in such a way that average noise level for the full patch corresponds to that used in the homogeneous noise cases studied earlier.

In Fig. 22 we show the variances of the B-mode spectra computed using full (left panel) and only well-observed part of the patch (right panel). The gain with respect to the noise variance largely compensates for the loss in terms of sample variance due to removal of the very noisy pixels close to the edge of the full survey area. This is especially apparent for $\ell < 150$. When analyzing the full sky survey, long apodization lengths are preferred as underlined by the red crosses displaying the minimal variance achieved for each bin. This is because a long apodization length allows us better to mitigate the influence of the noisiest pixels at the edge and thus to lower the noise variance. This also explains why small apodization lengths are suboptimal for the entire range of considered ℓ-modes. However, once the noisiest pixels are removed (top-right panel), the situation changes drastically. Moderate apodizations are preferred at large angular scales (4 degrees for $\ell < 100$). This value of the optimized apodization length is the same as for the homogeneous case. Nevertheless, at small angular scales, the behavior of the variance with respect to the apodization length for the inhomogeneous noise differs from the homogeneous noise case. For the latter, small, ~ 2 degrees, apodizations are preferred at small angular scales whereas long apodizations should be used at small scales, ~ 8 degrees for $\ell > 800$, for the former. Such a modification is due to the noise variance (we remind that with analytic windows, the E-to-B leakage is sufficently lowered below the noise level). In this part of the spectrum, windowing should converge towards inverse-noise weighting for the estimation to be close to optimal. In the homogeneous noise case, this requires to lower as much as possible the apodization length. In the inhomogeneous case, however, this means that windowing should be as close as possible to the actual noise distribution. For our realistic example, the noise distribution as shown on Fig. 21, exhibits a relatively high apodization length, which translates into a large apodization needed for the window to be optimal.

In a more general case the question if any given analytic window can produce the results close to the optimal is more complex and will depend on both the window and the noise distribution. In fact, at high ℓ, the near optimality can be reached only if the adopted window function approximates the inverse noise weighting.
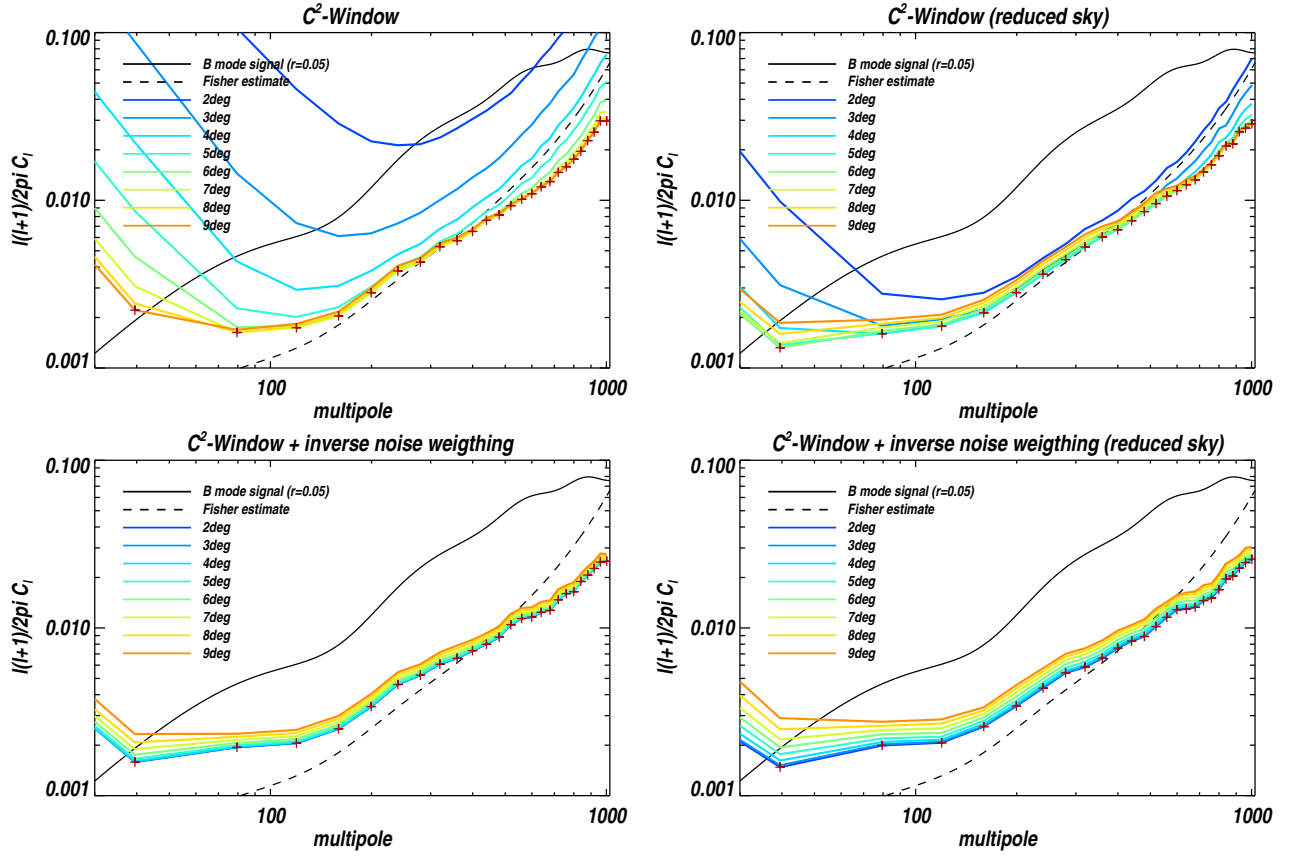
FIG. 22: *Upper panels:* Variance of the estimated B-mode power spectrum computed using C^2-windows with different apodization length (color curves) for the full observed (left panel) and the well-observed (right panel) sky area. The black-dashed curves show the Fisher estimate of the variance calculated assuming the average noise level (see text) and are only shown as a reference here. The black solid curve represents the input B-mode power spectrum. The red crosses indicate the lowest variance found for each bin, thus defining the optimized apodization length selected via the MC optimization procedure. For a given apodization length, removing the noisiest pixel significantly reduces the noise. Nevertheless, except for the first two bins the optimized variance is nearly the same for the two considered sky patches. *Lower panels:* Same as above but using a window function given as a product of the analytic C^2-window and the inverse square-noise variance. This extra noise weighting largely alleviates the problem of the noisiest pixels resulting in similar variance estimates for the full and reduced observed sky patches. Moreover the optimized apodization length, marked with red crosses, is independent of the bin number.

Our discussion and results derived above show that the analytic windows introduced in this paper can indeed be appropriate for noise patterns similar to EBEX. A more general approach, applicable to other noise inhomogeneity patterns is also however possible. We first note that multiplying any of the analytic apodization studied here by some smooth function of the sky position will result in a proper apodization, which suppresses the E/B leakage. We emphasize that this observation is indeed strictly true only for nonpixelized skies and as usual a care has to be exercised whenever pixelization is introduced. In general, for pixelized skies the smooth function needs also to be sufficiently flat at the boundary not to lead to some uncontrolled level of leakage. In the case at hand we derive the new spin-0 window by multiplying the analytic windows by a pixel-dependent weight corresponding to the inverse squared noise weighting, $\propto 1/\sigma^2(p)$, where, $\sigma^2(p)$ is the estimated noise variance at pixel p. We then numerically compute the spin-1 and spin-2 components from it. This allows us to mimic inverse-noise weighting while fulfilling the proper boundary constraints. The effect of such inverse-noise weighting is displayed in the lower panels of Fig. 22 for the full survey (left panel) and its well-observed part (right panel), where colors refer to the value of the apodization length of W_0. The results show that using this inverse-noise weighting significanlty reduces the total variance, particularly for small apodization length. This explains the origin of the extra-variance at low-ℓ end seen in the upper panels of Fig. 22 relating it to the noise of the noisy pixels present at the map boundary.

Interestingly whatever survey we use (full or reduced), the lowest variance is always reached for a small apodization length ($\delta_c = 2$ degrees) apparently alleviating the need of going through the optimization procedure of the apodization length. We note however that this fact is
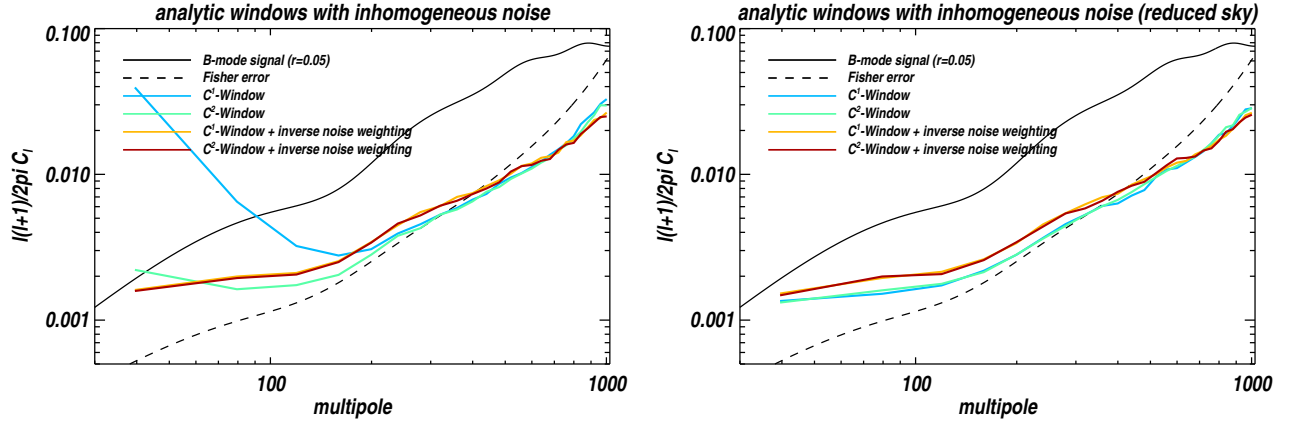
FIG. 23: Variance of the estimated B-mode cross-spectra using the C^1- and C^2-windows as well as their inverse-noise weighted extensions for the full observed (left panel) and well-observed sky (right panel) areas. The black-dashed curves show the Fisher estimate of the variance derived for the average noise level and the black solid curve depicts the input B-mode power spectrum. If the full observed map is to be analyzed, adding inverse-noise weighting suppresses the variance below the estimated power spectrum level at the low-ℓ regime. With the noisiest pixels removed, adding inverse-noise weighting does not change much except at large angular scale where it results in some excess variance, when compared with the simple analytic windowing.

related to the specific noise distribution considered here and will be applicable in general only to the cases with a sufficient, noise-induced apodization present at the patch edges, which, as shown here, are however of a practical interest. At the high-ℓ part, this is because the near optimality is reached thanks to the inverse-noise weighting and therefore the small apodization length of the analytic part of the window is preferred, as it affects the total window shape to the least extent. At the low-ℓ part of the spectrum, without the extra noise-weighting an apodization length of 4 degrees was needed for the analytic weighting alone. The effective, noise-induced apodization is larger than that and the apodization length for the analytic part of the window can be therefore smaller as the latter is needed only to fulfill the boundary conditions and reduce the E/B leakage.

Somewhat counter-intuitively the numerically derived estimates of the power spectrum uncertainties are seen in Fig. 22 to be superior to the forecasts based on the Fisher formalism. This is due to the fact that the latter are based on the average noise level computed assuming the same number of observed samples as in the actual case, but now homogeneously distributed over the entire observed sky area, and the numerical calculations benefit, thanks to either noise-weighting or appropriate ℓ-bin dependent apodization, from the very low noise core of the patch, which dominates the constraints at the high-ℓ end of the spectrum.

A comparison of the best estimates of the spectral variance derived with different analytic windows is shown in Fig. 23. It demonstrates the importance of the down-weighting of the very noisy pixels for the full patch analysis (left panel), which is needed to achieve a variance level lower than the input B-mode at large angular scales. This effect is particularly prominent for the C^1-window, which by construction is less smooth towards the patch

boundary than the C^2-window, and adding inverse noise weighting is here mandatory. If the noisiest pixels are removed from the analysis (right panel), using inverse-noise weighting performs the same as analytic windows alone except for $\ell < 200$, where inverse-noise weighting is somewhat suboptimal as compared to analytic windowing alone. This effect is analogous to the usual loss of precision in the signal-dominated regime also seen in the standard pseudo power spectrum cases. However, as noted previously, the results derived with the explicit inverse-noise weighting have not required any MC-based optimization in the case studied here. Therefore though suboptimal for $\ell < 200$ they can certainly be used to provide a first and reliable estimate of the B-mode power spectrum, at least in the cases whenever the noise distribution complies with the assumptions as already spelled out earlier.

In Fig. 24 we show the variances derived using the numerically optimized windows and compare them against those obtained with the help of the analytic ones discussed already before. Only the reduced, well-observed sky area has been used here, due to the numerical convergence problems experienced by the PCG procedure as mentioned before. The apodization length of the analytic windows (which include no noise-weighting here) has been optimized in each bin in order to minimize the variance. Clearly, the windows optimized numerically in the pixel-domain display very similar levels of the uncertainty on the recovered power spectrum level as derived with the analytic apodizations. The harmonic space optimized window shows however an excess of variance in the entire range of considered ℓ-modes. This is because it has been calculated using the average noise level, which in all bins has incorporated all the pixels, including those close to the patch boundary, therefore it does not benefit from the low-variance of the pixels in the very center of
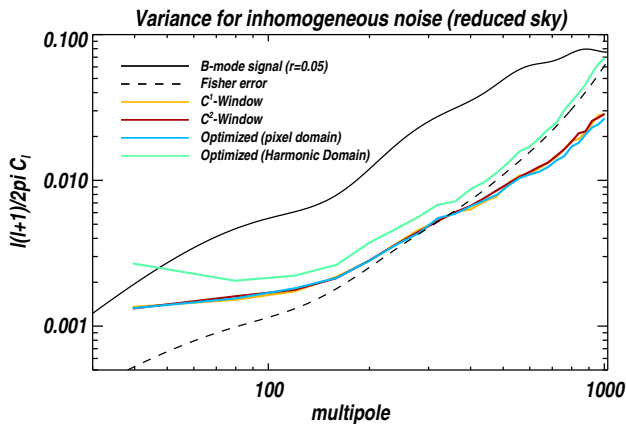
FIG. 24: Variance of the estimated B-mode power spectrum for the realistic study case considered here. Four different windows have been used : C^1-window (yellow curve), C^2-window (red curve), optimized window computed in the harmonic domain (green curve) and the optimized window computed in the pixel-domain (blue curve). The apodization length of the analytic windows has been optimized in each bin. The black-dashed curve stands for the Fisher estimates of the variance and the black-solid one for the input theoretical B-mode. The Fisher estimate has been calculated assuming the average noise level. All the windows show comparable performance with an exception of the harmonic domain optimized one. This departure is due to the homogeneous approximation of the noise making the effective apodization length of the harmonic optimized window suboptimal.

the patch. At the high-ℓ end the variance obtained with the harmonic domain window follows closely the Fisher matrix estimate, which has been derived using the same approximation for the adopted noise level.

V.   CONCLUSIONS

We have implemented the pure pseudo-spectrum formalism of [15, 16] extending it to incorporate the cross-spectrum approach and discussed practical issues related to its implementation. In particular, we have presented a quick, efficient way to compute pure multipoles and their respective mode mixing kernels utilizing harmonic domain derivatives, and demonstrated its overall consistency. We have also considered a number of apodization proposals. Those have included analytical, semianalytical and numerically-derived functions. We have studied the performance of the implemented formalism from the point of view of a level of the residual E/B leakage as well as final variance of the estimated spectra, and investigated their dependences on a number of factors such as an observed sky geometry, window choice, pixelization type and pixel size. We have considered the role of the sky priors in the apodization optimization procedures, as well as proposed an iterative, prior-free approach to the B-mode power spectrum estimation.

Our results support the assertion of [15, 16] and show that the pure pseudo spectrum approach not only can successfully resolve the E/B leakage problem but, if properly applied, it also allows to bring down the total variance of the estimated power spectrum to levels comparable with those derived from the optimal approaches. And it can do that at the significantly lower computational costs. In fact, we usually find that the power spectrum variance derived with the pure approach is within a factor of 2 away from the reference variance based on the Fisher matrix approach, which neglects the E/B leakage and assumes no apodization (but a binary mask).

We have shown that the analytic apodizations, in particular the C^2-window proposed here, can deliver a competitive performance, particularly if optimized using the MC simulations, at the low calculational cost. The applicability of the analytic windows can be greatly extended by combining them with the inverse square-noise weighting. Such weighting may sometimes produce somewhat suboptimal results at the intermediate and low end of the ℓ-range, where it can/should be replaced by the unweighted analytic or numerically optimized windows. We have however found that in some cases of practical interest the noise weighted analytic windows may not require any optimization procedure. They therefore provide a suitable first guess useful in particular in a preliminary quick-look analysis.

In the cases with a nearly homogeneous or mildly inhomogeneous noise distribution, we have derived a semianalytic, quick-to-compute harmonic domain window, which though requires some prior information about the sky signal, provides yet another cheap and efficient alternative to be used in the pure spectrum calculation. However, in the highly inhomogeneous cases such a procedure may be suboptimal and further work is needed to bring its performance in-line with the other methods discussed here.

The numerical window optimization as proposed in [16] have been shown to provide an efficient framework for the optimized window computation applicable to a wide variety of the problems. It can however be numerically prohibitive and iterative solvers needed to overcome this problem may suffer due to lack of convergence. We have found that a particularly common case is whenever a high dynamic range of expected noise levels is present. In such circumstances a conservative tresholding procedure may have to be applied in practice to avoid such difficulties.

We note that although these conclusions have been derived from the cross-spectrum analysis they should be equally applicable to autospectra cases.

In this work we have not considered issues related to bin-bin correlations of the power spectrum estimates. This is an important topic for the further applications of the results of the method discussed here, in particular for the cosmological parameter estimation, and will be treated in the forthcoming work.

This paper focuses only on the statistical uncertainty incurred due to either the noise and/or sampling variance. Other potential sources of error, for instance, re-

sulting from a foreground component separation and/or systematic effect treatment are not included. Clearly, such effects will have to be incorporated in a final error budget of B-mode power spectrum estimated from any realistic CMB polarization data. The results presented here set a stage for an analysis of those issues in future work.

In this work, we have not discussed the role and a potential impact of pixels missing from the main body of the observed map. Such "holes" in the maps could result from, for example, a point-sources removal or masking. We note that their presence does not lead to any fundamental problems in the pure formalism, as has been shown in the case of the pixel-domain optimized windows in [16]. However, the practical issues exist. In particular in the cases with the analytic windowing, whenever the missing pixels are closer to each other and/or observed patch edge than the considered apodization length a special care must be taken to avoid windows discontinuities (Sect. IV A 4). Such effects may cause some additional E/B leakage, however they are unlikely to lead to any appreciable increase of the total variance of the recovered B-mode power spectrum. A more detailed study of the effect of holes in B-mode reconstruction will be presented in a future work.

The results derived in this paper were obtained with the help of a publicly available software package S^{2}HAT [30]. The set of MPI-parallel routines permitting quick computation of the pure multipoles as well as spin-weighted apodizations given an input spin-0 window, developed in the course of this work, can be downloaded from [37].

Acknowledgments

APPENDIX A: MIXING KERNEL

The mixing kernel is implicitly defined by Eq. (12). It can be explicitly derived in the harmonic space by expressing the pseudo-C_ℓ's, averaged over CMB realizations, as functions of the CMB power spectrum. The derivation is very similar to the one described in [11]. We first recall that the pure pseudo-$a_{\ell m}$'s are given by the following expression:

$$
_E\tilde{a}_{\ell m} = E_{2,\ell m} + 2\frac{N_{\ell,1}}{N_{\ell,2}}E_{1,\ell m} + \frac{1}{N_{\ell,2}}E_{0,\ell m}, \quad \text{(A1)}
$$

$$
_B\tilde{a}_{\ell m} = B_{2,\ell m} + 2\frac{N_{\ell,1}}{N_{\ell,2}}B_{1,\ell m} + \frac{1}{N_{\ell,2}}B_{0,\ell m}, \quad \text{(A2)}
$$

where $N_{\ell,s} \equiv \sqrt{(\ell+s)!/(\ell-s)!}$. Employing now harmonic space decompositions for the polarization field $Q + iU$ and the different spin-s window functions, we can show that,

$$
\begin{aligned}
E_{s,\ell m} =& \frac{(-1)^m}{4}\sum_{\ell'm'}\sum_{\ell''m''}N_{\ell'',(2-s)}\sqrt{\frac{(2\ell+1)(2\ell'+1)(2\ell''+1)}{4\pi}}w_{\ell''m''}\begin{pmatrix}\ell & \ell' & \ell'' \\ -m & m' & m''\end{pmatrix} \\
&\times\left\{\left[\begin{pmatrix}\ell & \ell' & \ell'' \\ -s & 2 & -2+s\end{pmatrix} + \begin{pmatrix}\ell & \ell' & \ell'' \\ s & -2 & 2-s\end{pmatrix}\right]a^E_{\ell'm'} + i\left[\begin{pmatrix}\ell & \ell' & \ell'' \\ -s & 2 & -2+s\end{pmatrix} - \begin{pmatrix}\ell & \ell' & \ell'' \\ s & -2 & 2-s\end{pmatrix}\right]a^B_{\ell'm'}\right\}, \\
B_{s,\ell m} =& \frac{(-1)^m}{4}\sum_{\ell'm'}\sum_{\ell''m''}N_{\ell'',2-s}\sqrt{\frac{(2\ell+1)(2\ell'+1)(2\ell''+1)}{4\pi}}w_{\ell''m''}\begin{pmatrix}\ell & \ell' & \ell'' \\ -m & m' & m''\end{pmatrix} \\
&\times\left\{\left[\begin{pmatrix}\ell & \ell' & \ell'' \\ -s & 2 & -2+s\end{pmatrix} + \begin{pmatrix}\ell & \ell' & \ell'' \\ s & -2 & 2-s\end{pmatrix}\right]a^B_{\ell'm'} - i\left[\begin{pmatrix}\ell & \ell' & \ell'' \\ -s & 2 & -2+s\end{pmatrix} - \begin{pmatrix}\ell & \ell' & \ell'' \\ s & -2 & 2-s\end{pmatrix}\right]a^E_{\ell'm'}\right\}.
\end{aligned}
$$

The above expressions can be then inserted into Eqs. (A1) and (A2), which we can use to define the pure pseudo-C_ℓ's. Taking then the average over CMB realizations and making use of the orthogonality relations,

$$
(2\ell''+1)\sum_{mm'}\begin{pmatrix}\ell & \ell' & \ell'' \\ -m & m' & m''\end{pmatrix}\begin{pmatrix}\ell & \ell' & \ell''' \\ -m & m' & m'''\end{pmatrix} = \delta_{\ell''\ell'''}\delta_{m''m'''},
$$

we arrive at,

$$\begin{pmatrix} \tilde{C}^E_\ell \\ \tilde{C}^B_\ell \end{pmatrix} = \sum_{\ell'} \begin{pmatrix} M^{diag}_{\ell\ell'} & M^{off}_{\ell\ell'} \\ M^{off}_{\ell\ell'} & M^{diag}_{\ell\ell'} \end{pmatrix} \begin{pmatrix} C^E_{\ell'} \\ C^B_{\ell'} \end{pmatrix}. \tag{A3}$$

On defining the power spectrum of the spin-0 window function by

$$\mathcal{W}_{\ell''} = \frac{1}{2\ell''+1} \sum_{m''} w_{\ell''m''} w^\dagger_{\ell''m''},$$

and introducing the following notation

$$J^\pm_s(\ell,\ell',\ell'') = \begin{pmatrix} \ell & \ell' & \ell'' \\ -2+s & 2 & -s \end{pmatrix} \pm \begin{pmatrix} \ell & \ell' & \ell'' \\ 2-s & -2 & s \end{pmatrix},$$

we can express the diagonal and off-diagonal blocks (in the E/B subspaces) of the mixing matrix as:

$$M^{diag}_{\ell\ell'} = \frac{2\ell'+1}{16\pi} \sum_{\ell''} (2\ell''+1)\mathcal{W}_{\ell''} \left[J^+_0(\ell,\ell',\ell'') + 2\sqrt{\frac{(\ell+1)!(\ell-2)!(\ell''+1)!}{(\ell-1)!(\ell+2)!(\ell''-1)!}} J^+_1(\ell,\ell',\ell'') \right. \tag{A4}$$
$$\left. + \sqrt{\frac{(\ell-2)!(\ell''+2)!}{(\ell+2)!(\ell''-2)!}} J^+_2(\ell,\ell',\ell'') \right]^2,$$

$$M^{off}_{\ell\ell'} = \frac{2\ell'+1}{16\pi} \sum_{\ell''} (2\ell''+1)\mathcal{W}_{\ell''} \left[J^-_0(\ell,\ell',\ell'') + 2\sqrt{\frac{(\ell+1)!(\ell-2)!(\ell''+1)!}{(\ell-1)!(\ell+2)!(\ell''-1)!}} J^-_1(\ell,\ell',\ell'') \right. \tag{A5}$$
$$\left. + \sqrt{\frac{(\ell-2)!(\ell''+2)!}{(\ell+2)!(\ell''-2)!}} J^-_2(\ell,\ell',\ell'') \right]^2.$$

We note that in the standard pseudo-spectrum approach, the mixing kernels are obtained by removing the $J^\pm_1$ and $J^\pm_2$ terms in the above expressions as the latter encode the contribution of the counterterms.

If the window function is apodized up to its first derivative, the off-diagonal block of the mixing matrix should be equal to zero. We have checked numerically that this is indeed the case, when the above formulae are applied.

For the binned power spectrum estimation, the mixing matrix as implicitly defined in Eq. (12), is given by,

$$M^{XX'}_{\alpha\alpha'} = \sum_{\ell\in\alpha} \sum_{\ell'\in\alpha'} \frac{\ell(\ell+1)}{\ell'(\ell'+1)\Delta\ell} M^{diag}_{\ell\ell'} \delta_{XX'}. \tag{A6}$$

In the above formulae we have assumed that the derivative constraints between the three spin-s windows are exactly fulfilled. However, in practice they may be only approximately satisfied and the mixing kernel should be therefore expressed not only as a function of W, but also of W_1 and W_2. If such extensions are incorporated, the mixing kernel reads,

$$M^{diag}_{\ell\ell'} = \frac{2\ell'+1}{16\pi} \sum_{\ell''m''} \left| w_{0,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix} + w_{0,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 2 & -2 & 0 \end{pmatrix} \right. \tag{A7}$$
$$+ 2\sqrt{\frac{(\ell+1)!(\ell-2)!}{(\ell-1)!(\ell+2)!}} \left[-w_{-1,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ -1 & 2 & -1 \end{pmatrix} + w_{1,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 1 & -2 & 1 \end{pmatrix} \right]$$
$$\left. + \sqrt{\frac{(\ell-2)!}{(\ell+2)!}} \left[w_{-2,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 2 & -2 \end{pmatrix} + w_{2,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & -2 & 2 \end{pmatrix} \right] \right|^2$$

$$M_{\ell\ell'}^{off} = \frac{2\ell'+1}{16\pi} \sum_{\ell''m''} \left| w_{0,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ -2 & 2 & 0 \end{pmatrix} - w_{0,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 2 & -2 & 0 \end{pmatrix} \right.$$

$$+2\sqrt{\frac{(\ell+1)!(\ell-2)!}{(\ell-1)!(\ell+2)!}} \left[-w_{-1,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ -1 & 2 & -1 \end{pmatrix} - w_{1,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 1 & -2 & 1 \end{pmatrix} \right]$$

$$\left. +\sqrt{\frac{(\ell-2)!}{(\ell+2)!}} \left[w_{-2,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 2 & -2 \end{pmatrix} - w_{2,\ell''m''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & -2 & 2 \end{pmatrix} \right] \right|^2, \tag{A8}$$

where

$$w_{\pm s,\ell''m''} = \int_{4\pi} W_{\pm s}(\Omega) {}_{\pm s}Y^\dagger_{\ell''m''}(\Omega) d^2\Omega. \tag{A9}$$

These last expressions can be rewritten using the E and B decomposition of the spin-weighted window functions, defined as,

$$w_{0,\ell m}^{(E)} = -w_{0,\ell m}, \tag{A10}$$

$$w_{s,\ell m}^{(E)} = -\frac{1}{2}\left(w_{s,\ell m} + (-1)^s w_{-s,\ell m}\right), \tag{A11}$$

$$w_{s,\ell m}^{(B)} = \frac{i}{2}\left(w_{s,\ell m} - (-1)^s w_{-s,\ell m}\right), \tag{A12}$$

with $s = 1$ or 2. Using this decomposition in the expression for the mixing kernel leads to,

$$M_{\ell\ell'}^{diag} = \frac{2\ell'+1}{16\pi} \sum_{\ell''m''} \left| w_{0,\ell''m''}^{(E)} J_0^+ + 2\sqrt{\frac{(\ell+1)!(\ell-2)!}{(\ell-1)!(\ell+2)!}} w_{1,\ell''m''}^{(E)} J_1^+ + \sqrt{\frac{(\ell-2)!}{(\ell+2)!}} w_{2,\ell''m''}^{(E)} J_2^+ \right|^2$$

$$+\left| 2\sqrt{\frac{(\ell+1)!(\ell-2)!}{(\ell-1)!(\ell+2)!}} w_{1,\ell''m''}^{(B)} J_1^- + \sqrt{\frac{(\ell-2)!}{(\ell+2)!}} w_{2,\ell''m''}^{(B)} J_2^- \right|^2, \tag{A13}$$

$$M_{\ell\ell'}^{off} = \frac{2\ell'+1}{16\pi} \sum_{\ell''m''} \left| w_{0,\ell''m''}^{(E)} J_0^- + 2\sqrt{\frac{(\ell+1)!(\ell-2)!}{(\ell-1)!(\ell+2)!}} w_{1,\ell''m''}^{(E)} J_1^- + \sqrt{\frac{(\ell-2)!}{(\ell+2)!}} w_{2,\ell''m''}^{(E)} J_2^- \right|^2$$

$$+\left| 2\sqrt{\frac{(\ell+1)!(\ell-2)!}{(\ell-1)!(\ell+2)!}} w_{1,\ell''m''}^{(B)} J_1^+ + \sqrt{\frac{(\ell-2)!}{(\ell+2)!}} w_{2,\ell''m''}^{(B)} J_2^+ \right|^2. \tag{A14}$$

We point out that because the spin-0 window is real, it does not have any B component. As a consequence, when the derivative relation is completely fulfilled, the spin-1 and spin-2 windows have also a vanishing B part and the second line in Eqs. (A13) and (A14) are equal to zero. This is the case when using analytic windows or optimized windows computed in the harmonic domain. However, because the derivative relation is relaxed during the pixel implementation of the optimized windows, the B part of the two spin-weighted windows is not zero anymore in this special case, though it remains much smaller than their E part.

Equations (A13) and (A14) are directly applicable to the mixing kernel computations for both auto- and cross-spectra, but in the latter case only when the same apodization is adopted for all the maps. However two different apodizations are often required in the cross-spectrum case, for example,x because noise properties of

the two maps differ, and then the above equations have to be appropriately generalized. This can be done by representing each of the moduli present on the right hand sides of Eqs. (A13) and (A14) as a product of a respective expression, of which the modulus is taken, and its complex conjugate and subsequently using one of the two different apodizations for calculations of the expression and the other – its conjugate. For example, the square modulus

$$\left| 2\frac{N_{\ell,1}}{N_{\ell,2}} J_1^- w_{1,\ell''m''}^{(B)} + \frac{1}{N_{\ell,2}} J_2^- w_{2,\ell''m''}^{(B)} \right|^2$$

in Eq. (A13), is replaced by

$$\left(2\frac{N_{\ell,1}}{N_{\ell,2}} J_{1\ (0)}^- w_{1,\ell''m''}^{(B)} + \frac{1}{N_{\ell,2}} J_{2\ (0)}^- w_{2,\ell''m''}^{(B)} \right)^*$$

$$\times \left(2\frac{N_{\ell,1}}{N_{\ell,2}} J_{1\ (1)}^- w_{1,\ell''m''}^{(B)} + \frac{1}{N_{\ell,2}} J_{2\ (1)}^- w_{2,\ell''m''}^{(B)} \right),$$

where $_{(0)}w_{s,\ell''m''}^{(B)}$ and $_{(1)}w_{s,\ell''m''}^{(B)}$ are the window functions applied to the first and second map respectively.

In this paper we use Eqs. (A13) and (A14), or their generalization as described above, for the calculations of all the mixing kernels.

APPENDIX B: OPTIMIZED WINDOW FOR CROSS-SPECTRA

In this Appendix, we show how the optimized ansatz proposed in [16] in the context of the autospectra can be extended to incorporate the case of cross-spectra. Depending on the general context, different choices can be adopted for such an extension. We will present three options, focusing on the case of two maps, $\mathbf{d}^{(0)}$ and $\mathbf{d}^{(1)}$. In the following, i, j indices will run over pixels and A, B indices over maps.

The first and simplest choice is to optimize with respect to the auto pseudo-spectrum of each of the maps. The optimized windows are then directly given by the ansatz of [16] applied to each of the maps separately,

$$\sum_{j=1}^{N_{obs}} \mathbf{C}_{ij}^{(00)} \mathbf{P}_{ij}^{(\alpha)} W_i^{(0)} = 1, \quad (B1)$$

$$\sum_{j=1}^{N_{obs}} \mathbf{C}_{ij}^{(11)} \mathbf{P}_{ij}^{(\alpha)} W_i^{(1)} = 1, \quad (B2)$$

where $\mathbf{C}_{ij}^{AA} = \langle \mathbf{d}_i^A \mathbf{d}_j^A \rangle$ is the covariance matrix of the A-th map and $\mathbf{P}_{ij}^{(\alpha)}$ the geometrical matrix projecting in the ℓ, m space (see [16]).

The second option optimizes with respect to the autospectrum of the two maps combined together in the optimal way. In this case the optimal, quadratic estimator and the pseudo-spectrum one can be written as,

$$\mathcal{O}_\alpha = \sum_{i,j=1}^{N_{obs}} \sum_{A,B=0}^{1} \mathbf{d}_i^A \left(\mathbf{C}^{-1} \mathbf{P}^{(\alpha)} \mathbf{C}^{-1} \right)_{ij}^{AB} \mathbf{d}_j^B, \quad (B3)$$

$$\mathcal{C}_\alpha = \sum_{i,j=1}^{N_{obs}} \sum_{A,B=0}^{1} \mathbf{d}_i^A W_i^A \mathbf{P}_{ij}^{(\alpha)} W_j^B \mathbf{d}_j^B, \quad (B4)$$

where $\mathbf{C} \equiv \langle \mathbf{d}\mathbf{d}^t \rangle$ is the full covariance matrix of the data consisting of the two maps concatenated together, i.e.,

$$\mathbf{C} = \begin{pmatrix} \mathbf{S}_{ij} + \mathbf{N}_{ij}^{(00)} & \mathbf{S}_{ij} \\ \mathbf{S}_{ij} & \mathbf{S}_{ij} + \mathbf{N}_{ij}^{(11)} \end{pmatrix},$$

with $\mathbf{S}$ and $\mathbf{N}$ the signal and noise covariance matrices respectively. From the above definitions of the estimators, the formalism of [16] is easily generalized to get,

$$\sum_{i,j=0}^{N_{obs}} \sum_{B=0}^{1} \mathbf{C}_{ij}^{AB} \mathbf{P}_{ij}^{(\alpha)} W_j^B = 1, \quad (B5)$$

This is easily generalized to an arbitrary amount of maps by summing over the entire set of maps.

To compare the above optimized window to the one given by Eqs. (B1) & (B2), we assume the noise properties to be the same for the two maps. As a consequence, optimal windowing will be identical for each map. By plugging this into Eq. (B5), one can easily show that optimal windows coming from this last equation differs from optimal windows coming from Eqs. (B1) & (B2) by the way the noise is accounted for: with Eq. (B5), the noise part is divided by a factor 2 as compared to Eqs. (B1) & (B2). If N_{map} maps are to be analyzed, then the noise contribution would be divided by a factor N_{map} instead of 2.

The third and final option we consider consists of optimizing the weighting with respect to each cross-spectrum. For two maps, the optimal, quadratic and pseudo-sepctrum estimators read,

$$\mathcal{O}_\alpha = \sum_{i,j=1}^{N_{obs}} \mathbf{d}_i^{(0)} \left(\mathbf{C}^{-1} \mathbf{P}^{(\alpha)} \mathbf{C}^{-1} \right)_{ij}^{(01)} \mathbf{d}_j^{(1)}, \quad (B6)$$

$$\mathcal{C}_\alpha = \sum_{i,j=1}^{N_{obs}} \mathbf{d}_i^{(0)} W_i^{(01)} \mathbf{P}_{ij}^{(\alpha)} W_j^{(01)} \mathbf{d}_j^{(1)}. \quad (B7)$$

$$(B8)$$

The optimal, quadratic estimators can be recast as a function of the inverse of the covariance matrix $\mathbf{D} = \mathbf{C}^{-1}$,

$$\mathcal{O}_\alpha = \sum_{i,j=1}^{N_{obs}} \sum_{m,n=1}^{N_{obs}} \mathbf{d}_i^{(0)} \left(D^{(00)} + D^{(10)} \right)_{im}^{T} \mathbf{P}_{mn}^{(\alpha)} \left(D^{(01)} + D^{(11)} \right)_{nj} \mathbf{d}_j^{(1)}. \quad (B9)$$

The explicit expression of $\mathbf{D}$ are given by,

$$D^{(00)} = \left[\mathbf{C}^{(00)} - \mathbf{C}^{(01)} \mathbf{C}^{(11)^{-1}} \mathbf{C}^{(01)} \right]^{-1}, \quad (B10)$$

$$D^{(11)} = \left[\mathbf{C}^{(11)} - \mathbf{C}^{(01)} \mathbf{C}^{(00)^{-1}} \mathbf{C}^{(01)} \right]^{-1}, \quad (B11)$$

$$D^{(01)} = -\mathbf{C}^{(00)^{-1}} \mathbf{C}^{(01)} D^{(11)}, \quad (B12)$$

$$D^{(10)} = -\mathbf{C}^{(11)^{-1}} \mathbf{C}^{(01)} D^{(00)}. \quad (B13)$$

Following [16], the optimized window is obtained by minimizing the distance between the optimal, quadratic estimator and the cross pseudo-spectrum estimators with respect to $W^{(01)}$. This leads to the following linear sys-

tem,

$$\sum_{i=1}^{N_{obs}} \mathbf{C}_{ik}^{(01)} \mathbf{P}_{ik}^{(\alpha)} W_i^{(01)} = \tag{B14}$$

$$\frac{1}{2} \sum_{i,j=1}^{N_{obs}} \mathbf{C}_{jk}^{(01)} \mathbf{P}_{ik}^{(\alpha)} \left(D_{ij}^{(00)} + D_{ij}^{(10)} + D_{ij}^{(01)} + D_{ij}^{(11)} \right).$$

This last approach is probably the more appropriate as the resulting windows are optimized for each cross-spectra. However, due to the rather complicated right hand side in Eq. (B14), it is also more numerically involved than the two first approaches.

APPENDIX C: HARMONIC COMPUTATION OF OPTIMIZED WINDOWS

In this Appendix, we describe the major steps to compute the optimized windows in the case of white and homogeneous noise. As such noise is completely described by its power spectrum, we perform all the computations in the harmonic space. We start from Eqs. (18) & (19) and thus choose to use the χ-fields rather than the spin-weighted approach in parlance of [16] . Consequently, we will optimize only the spin-0 window and compute the spin-1 and spin-2 windows as its numerical derivatives. We remind that the χ^B field is a scalar field defined as [16]

$$\chi^B \equiv \sqrt{\frac{(\ell+2)!}{(\ell-2)!}} \mathbf{D}_2^B \cdot \mathbf{P}, \tag{C1}$$

where the ℓ-dependent prefactor compensates for the ℓ-dependant normalization of the $\mathbf{D}_2^B$ operator as defined in Eq. (9). The harmonic decomposition of the χ^B field reads

$$\chi^B = \sum_{\ell,m} \sqrt{\frac{(\ell+2)!}{(\ell-2)!}} a_{\ell m}^B Y_{\ell m}, \tag{C2}$$

where $a_{\ell m}^B$ denotes a type B multipole of the polarization vector $\mathbf{P}$.

Following the variational interpretation as proposed in [16], the optimized windows are the one which minimize the total aliased power in the pseudo-C_ℓ's. The B pseudo-spectrum averaged over CMB realization reads then,

$$\left\langle \tilde{C}_\ell \right\rangle = \sum_{\ell''m''} C_{\ell,\ell''}^\chi w_{0,\ell''m''} w_{0,\ell''m''}^*, \tag{C3}$$

where,

$$C_{\ell,\ell''}^\chi = \sum_{\ell'} \frac{2\ell'+1}{4\pi} \frac{(\ell-2)!(\ell'+2)!}{(\ell+2)!(\ell'-2)!} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2$$
$$\times \left(C_{\ell'}^B B_{\ell'}^2 + \sigma^2 \right), \tag{C4}$$

and $w_{0,\ell m}$ stands for the harmonic representation of the spin-0 window function. This window has to be apodized up to its first derivative to make Eqs. (18) & (19) and Eqs. (16) and (17) (i.e., the χ-fields and the spin-weighted approaches) equivalent.

As a consequence, we have to minimize $\left\langle \tilde{C}_\ell \right\rangle$ with respect to $w_{0,\ell m}$ under the three constraints. First of all, the spin-0 window has to be normalized within the observed region,

$$\sum_{\ell m} w_{0,\ell m} \mathcal{M}_{\ell m}^\dagger = 1, \tag{C5}$$

where $\mathcal{M}_{\ell m}$ is the harmonic representation of the mask. (The normalization constant can be set to unity without loss of generality.) Second, the spin-0 window and its first derivative, the spin-1 window, have to vanish at the contour of the observed region, denoted $\mathcal{P}$ in the following. This translates into $2N_c$ constrains, where N_c is the number of pixels on the contour

$$W(i) = 0 \text{ and } W_1(i) = 0, \quad \text{for all } i \in \mathcal{P}. \tag{C6}$$

With such a large number of external constraints, the standard Lagrange multiplier technique used to derive the optimized spin-0 window function becomes hard to solve. The complexity of the problem can be reduced by assuming that the integral of the spin-0 and the spin-1 windows on the contour have to vanish, leading to only two additional constraints

$$\sum_{\ell m} w_{0,\ell m} \mathcal{P}_{\ell m}^\dagger = 0, \tag{C7}$$

$$\sum_{\ell m} \sqrt{\ell(\ell+1)} w_{0,\ell m} \mathcal{P}_{\ell m}^\dagger = 0. \tag{C8}$$

Here $\mathcal{P}_{\ell m}$ is the harmonic representation of the contour. The second equation above corresponds to the spin-1 constraint as $w_{1,\ell m} = \sqrt{\ell(\ell+1)} w_{0,\ell m}$. The optimized spin-0 window can be then derived by applying the standard Lagrange multiplier techniques, resulting in,

$$w_{0,\ell''m''} = \frac{1}{C_{\ell,\ell''}^\chi} \left(\lambda \mathcal{M}_{\ell''m''} + \mu \mathcal{P}_{\ell''m''} \right). \tag{C9}$$

The two constants, λ and μ, are derived by plugging the above solution into the constraints. Defining the follow-

ing total power,

$$N_{A,0} = \sum_{\ell''m''} \frac{1}{C^\chi_{\ell,\ell''}} |\mathcal{M}_{\ell''m''}|^2, \tag{C10}$$

$$N_{A,1} = \sum_{\ell''m''} \frac{1}{C^\chi_{\ell,\ell''}} |\mathcal{P}_{\ell''m''}|^2, \tag{C11}$$

$$N_{A,2} = \sum_{\ell''m''} \frac{\ell(\ell+1)}{C^\chi_{\ell,\ell''}} |\mathcal{P}_{\ell''m''}|^2, \tag{C12}$$

$$N_{X,0} = \sum_{\ell''m''} \frac{1}{C^\chi_{\ell,\ell''}} \left|\mathcal{M}_{\ell''m''}\mathcal{P}^\dagger_{\ell''m''}\right|, \tag{C13}$$

$$N_{X,1} = \sum_{\ell''m''} \frac{\sqrt{\ell(\ell+1)}}{C^\chi_{\ell,\ell''}} \left|\mathcal{M}_{\ell''m''}\mathcal{P}^\dagger_{\ell''m''}\right|, \tag{C14}$$

$$N_{X,2} = \sum_{\ell''m''} \frac{\sqrt{\ell(\ell+1)}}{C^\chi_{\ell,\ell''}} |\mathcal{P}_{\ell''m''}|^2, \tag{C15}$$

the two constants read,

$$
\begin{aligned}
\lambda^{-1} \equiv\ & N_{A,0} \\
& - \frac{N^2_{X,0}N_{A,2} + N^2_{X,1}N_{A,1} - 2N_{X,0}N^2_{X,2}}{N_{A,1}N_{A,2} - N^2_{X,2}},
\end{aligned}
\tag{C16}
$$

$$
\begin{aligned}
\mu \equiv\ & -\frac{\lambda}{N_{A,1}N_{A,2} - N^2_{X,2}} \left[(N_{X,0}N_{A,2} - N_{X,1}N_{X,2}) \right.\\
& \left. - \sqrt{\ell(\ell+1)}\,(N_{X,0}N_{X,2} - N_{X,1}N_{A,1}) \right].
\end{aligned}
\tag{C17}
$$

In this derivation the boundary constraints have been partially relaxed, making the window only partially optimize. The numerical experiments, however, show that the resulting power spectra uncertainties are at the same level as the ones obtained using the exact, pixel-domain computation of the optimized window (see Fig. 18). This demonstrates that the proposed boundary constraint simplification does not compromise the performances of this fast computed, optimized window.

We point out that the above approach is general and does not apply only to the B-mode fields. For instance, we can use it to calculate optimized windows for a temperature (scalar) map. It suffices then to replace

$$\frac{(\ell-2)!(\ell'+2)!}{(\ell+2)!(\ell'-2)!}\left(C^B_{\ell'}B^2_{\ell'} + \sigma^2\right)$$

in Eq. (C4) by the total (signal plus noise) power spectra, i.e., $C^T_{\ell'}B^2_{\ell'} + \sigma^2$, of the map. Moreover, we can straightforwardly relax some of the boundary requirements. For example, if no boundary conditions are to be imposed, we just need to replace Eqs. (C16)-(C17) with,

$$\lambda = N^{-1}_{A,0} \tag{C18}$$

$$\mu = 0, \tag{C19}$$

while in a case when only the Dirichlet boundary condition is to be imposed, the two constants are given by:

$$\lambda = \frac{N_{A,1}}{N_{A,0}N_{A,1} - N^2_{X,0}} \tag{C20}$$

$$\mu = \frac{N_{X,0}}{N^2_{X,0} - N_{A,0}N_{A,1}}. \tag{C21}$$

[1] BICEP: http://www.astro.caltech.edu/~lgg/bicep_front.htm
[2] Planck: http://www.esa.int/esaSC/120398_index_0_m.html
[3] PolarBeaR: http://bolo.berkeley.edu/polarbear/science/science.html
[4] QUIET: http://quiet.uchicago.edu/
[5] Clover: http://www-astro.physics.ox.ac.uk/research/expcosmology/groupclover.html
[6] EBEX: http://groups.physics.umn.edu/cosmology/ebex/
[7] Spider: http://www.astro.caltech.edu/~lgg/spider_front.htm
[8] BRAIN: http://www.apc.univ-paris7.fr/APC_CS/Experiences/Brain/
[9] B-pol: http://www.b-pol.org/index.php
[10] M.G. Hauser, & P.J.E. Peebles, Astrophys. J., **185**, 757, (1973)
[11] E. Hivon, K.M. Górski, C.B. Netterfield, B.P. Crill, S. Prunet, & F. Hansen, Astrophys. J., **567**, 2, (2002)
[12] G. Hinshaw, et al., Astrophys.J.Supp., **148**,135, (2003)
[13] E.F. Bunn, M. Zaldarriaga, M. Tegmark, & A. de Oliveira-Costa, Phys. Rev. D **67**, 023501 (2003)
[14] G. Chon, A. Challinor, S. Prunet, & E. Hivon, MNRAS, **350**, 914, (2004)
[15] K.M. Smith, Phys. Rev. D **74** 083002 (2006)
[16] K.M. Smith & M. Zaldarriaga, Phys. Rev. D **76**, 043001, (2007)
[17] M. Tegmark & A. de Oliveira-Costa, Phys. Rev. D **64**, 063001, (2001)
[18] D.L. Larson, H.K. Eriksen, B.D. Wandelt, K.M. Górski, G. Huey, J.B. Jewell & I.J. O'Dwyer, Astrophys. J., **656**, 653 (2007)
[19] J.B. Jewell, H.K. Eriksen, B.D. Wandelt, I.J. O'Dwyer, G. Huey, K.M. Górski, arXiv:0807.0624
[20] M. Zaldarriaga & U. Seljak, Phys. Rev. D. **55**, 1830, (1997)
[21] M. Zaldarriaga & U. Seljak, Phys. Rev. D. **58**, 023003, (1998)
[22] D.A. Varshalovich, A.N. Moskalev & V.K. Khersonskii, Quantum Theory of Angular Momentum, (World Scientific, 1988)
[23] M. Kamionkowski, A. Kosowsky, & A. Stebbins, Phys. Rev. D, **55**, 7368, (1997)
[24] A. Kogut, et al, Astrophys.J.Supp., **148**, 161, (2003)
[25] F. Hansen & K.M. Górski, MNRAS, **343**, 559, (2003)
[26] M. Tristram, J.F. Macias-Pérez, C. Renault & D. Santos, MNRAS, **358**, 833, (2005)
[27] G. Polenta, D. Marinucci, A. Balbi, P. de Bernardis, E. Hivon, S. Masi, P. Natoli & N. Vittorio, JCAP **0511**,

001, (2005)

[28] P. Oxley, *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **5543**, 320, (2004)

[29] A. H. Jaffe, M. Kamionkowski & L. Wang, Phys. Rev. D **61**, 083501, (2000)

[30] `http://www.apc.univ-paris7.fr/∼radek/s2hat.html`

[31] K.M. Górski, E. Hivon, A.J. Banday, B.D. Wandelt,, F.K. Hansen, M. Reinecke & M. Bartelmann, Astrophys. J., **622**, 759, (2005) `http://healpix.jpl.nasa.gov`

[32] A.G. Doroshkievich, et al., Inter. Journal of Mod. Phys. D, 14, 275 (2005)

[33] P. F. Muciaccia, P. Natoli & N. Vittorio, Astrophys.J., 488, L63, (1997)

[34] A. Lewis, Phys. Rev. D., 71, 083008, (2005)

[35] G.H. Golub & C. F. van Loan, *Matrix operations*, (Johns Hopkins University Press, Baltimore, 1987)

[36] Dunkley, J., et al, `preprint arXiv:0803.0586`, (2008)

[37] `http://www.apc.univ-paris7.fr/∼radek/pmult.html`

[38] CMBfast: `http://www.cfa.harvard.edu/∼mzaldarr/ CMBFAST/cmbfast.html`

[39] CAMB: `http://camb.info/`

[40] `http://www.apc.univ-paris7.fr/∼betoule/ doku.php?id=fr:software`